Spring 7-1-2015

Bactericera Cockerelli (SULC) (hemiptera: triozidae): Nearly Complete Mitochondrial Genome, Single Mitochondrial Genes and Complete Mitochondrial Genome Comparison, and Insecticide Resistance Genes of Potato Psyllid Populations from North America

Amalia Rosa Lopez Montiel

BACTERICERA COCKERELLI (SULC) (HEMIPTERA: TRIOZIDAE): NEARLY COMPLETE MITOCHONDRIAL GENOME, SINGLE MITOCHONDRIAL GENES AND COMPLETE MITHOCONDRIAL GENOME COMPARISON, AND INSECTICIDE RESISTANCE GENES OF POTATO PSYLLID POPULATIONS FROM NORTH AMERICA

by

AMALIA ROSA LOPEZ MONTIEL

A Thesis submitted in partial fulfillment
of the requirements for the degree of
Masters of Science
Department of Biology

College of Arts and Sciences

The University of Texas at Tyler
July 2015

The University of Texas at Tyler

Tyler, Texas

This is to certify that the Master's Thesis of

AMALIA ROSA LOPEZ MONTIEL

has been approved for the thesis requirement on

July 14, 2015

for the Master of Science in Biology degree

Approvals:

Thesis Chair: Blake Bextine, Ph.D.

Member: Josh Banta, Ph.D.

Member: John Placyk, Ph.D.

Chair, Department of Biology: Srini Kambhampati, Ph.D.

Dean, College of Arts and Sciences: Martin Slann, Ph.D.

Acknowledgments

I would like to thank Dr. Blake Bextine for helping me out to reach one more goal of my professional development. Thank you for trusting and not hesitating in giving me the opportunity of being part of your lab. I am truly grateful!

My thesis committee (Dr. Josh Banta and Dr. John Placyk) and UT Tyler faculty of the Biology department, thanks for your guidance, instruction, and patience. Thank you for teaching me more than just science.

Jessica Coleman, thanks for the TA advising and making the experience less painful. I believe I learned more than the students! Dr. Kate Hertweck, for your guidance and help when the bioinformatics world was against me; and Dr. Brent Bill for sharing your knowledge, even though I was skeptical, had a great time in cell bio!

Maqueta Humber, thanks for dealing with all my paperwork. International students can become a huge nuisance; and Brad Low, thanks for accepting last minute orders, definitely wouldn't have finished my research on time!

My fellow graduate students, each one of you added a different flavor to my "American experience". It was really nice to meet you all!

Bextine lab members: Chissa, Patrick, Megan, Marie, Bijay, Chris, Juan, Mikayla, Mackenzie, and Katrina, for being such an amazing people and always willing to help. Special thanks to Daymon for sharing his knowledge and guiding me in my research. Juan, for being such an awesome fairy science and Chris for your help and for pushing and encouraging me to finish my thesis.

Kylie Swisher and lab, Margarita Palmieri and others who contributed in providing the basic element of my research, the potato psyllids.

My Tyler family, the Lambeths: Shauna, Jim, Lauren, Julia, Kent and Kacie, I wouldn't have survived without you. Thank you for your care, making me feel home, and a member of your wonderful family. I really have no words to express how grateful I am!

My roommates, Lauren and Bri, for being not only friends but my sisters!

My Guatemalan friends, especially to Pris and Totto, who despite the distance are always there. Thanks for those calls and texts. Your friendship is truly a gift!

Finally, my parents, brothers and little sister, for your love, concern and unconditional support. You are my engine to keep on doing what I like.

For all, I am eternally grateful!

Table of Contents

List of Tables

List of Figures

Abstract

BACTERICERA COCKERELLI (SULC) (HEMIPTERA: TRIOZIDAE): NEARLY COMPLETE MITOCHONDRIAL GENOME, SINGLE MITOCHONDRIAL GENES AND COMPLETE MITHOCONDRIAL GENOME COMPARISON, AND INSECTICIDE RESISTANCE GENES OF POTATO PSYLLID POPULATIONS FROM NORTH AMERICA

Amalia Rosa Lopez Montiel

Thesis Chair: Blake Bextine, Ph. D.

The University of Texas at Tyler
July 2015

The potato psyllid, *Bactericera cockerelli* (Sulc.), is an economic pest of potatoes and other solanaceous crops, and the vector of the causal agent of zebra chip disease. The expanded distribution of the potato psyllid has resulted in the differentiation of biotypes (central and western), and haplotypes (southwestern and northwestern), according to analysis of the mitochondrial gene cytochrome oxidase I (COI). In this study, the nearly complete (part of the control region unsequenced) mitochondrial genome from potato psyllid populations occurring in North America: Texas, Nebraska, California, Washington, northwest (Washington) and southwest (Arizona), were sequenced using next generation sequencing technology. The raw sequences files were assembled using Geneious and annotated with the software MITOS, DOGMA, ARWEN and NCBI ORF finder tool. The resulting mitochondrial genomes were studied for genome composition, organization and phylogenetic analysis. The gene content of the potato psyllid mitochondrial genomes included: 2 ribosomal RNAs, the 22 tRNAs and 13 protein coding genes. Through analysis of the nearly complete mitochondrial genome, potato psyllid populations from the central biotype (Texas and Nebraska) and the southwestern haplotype were determined to be

most closely related. Divergent from this clade were the Washington and California populations (western biotype), and the northwestern haplotype seems to be a new population. Importantly, analysis from single mitochondrial genes (COI, COIII, and ATP6) inferred the same phylogeny as one based on the whole mt genome. Moreover, the insecticide resistance gene cytochrome P45O4G had variations among populations that may be useful for future population genetic and phylogenetic studies.

Chapter 1

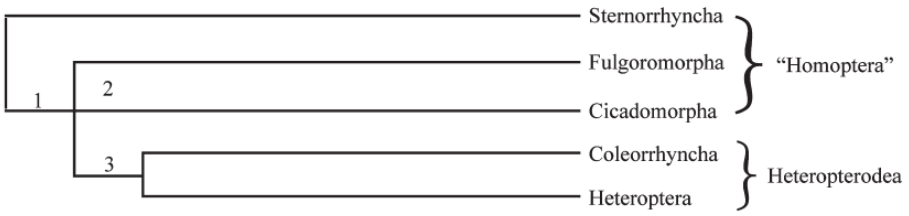
Literature review: Bactericera *cockerelli* (Sulc) (Hemiptera: Psylloidea: Triozidae)

Order Hemiptera

Hemiptera, an ancient lineage known to arise in the early Permian (Forero 2008), has 82,000 described species and is the largest group of nonholometablolous (incomplete metamorphosis) insect orders in the superorder Paraneoptera (Cui et al. 2013; Cryan and Urban 2012). The order comprises a wide diversity of morphological, biological adaptions, means of dispersal, and capability to transmit plant pathogens (Song et al. 2012; Forero 2008). To date, a consensus exists that this order includes the following major monophyletic clades: Sternorrhyncha, which is represented by scale insects, aphids, whiteflies, psyllids and others, and approximately 21 extant families; Heteroptera, with the true bugs and 54 extant families; Coleorrhyncha, also called the "moss bugs", comprised of one extant family; Fulgoromorpha, with the planthoppers and about 20 extant families; and Cicadomorpha, which comprises leafhoppers, treehoppers, spittlebugs, cicads and around 12 extant families (Cryan and Urban 2012). However, this order has been a topic of continuous debate among entomologists regarding its higher level of classification (Cryan and Urban 2012; Lee et al. 2009).

Before the use of "Hemiptera" as an order, Heteroptera and Homoptera were used to refer two independent orders, with Stenorrhyncha and Auchenorrhyncha as suborders of Homoptera (Cui et al. 203; Song et al. 2012; Lee et al. 2009). In addition, others considered Heteroptera and Homoptera as equivalently ranked suborders of Hemiptera (Cui et al. 2013; Song et al. 2012; Lee et al. 2009; Forero 2008; von Dohlen and Moran 1995). Since 1960 to the present, the concept of Hemiptera has been accepted (Cui et al. 2013), but different phylogenies between the higher-level

hemipteran lineages have been proposed. Cui et al. (2013), summarized the different views regarding the phylogeny of Hemiptera (Figure 1.1). In all hemipteran phylogeny views, Homoptera, whether is it considered as an order or as a suborder consists of Sternorryncha and Auchenorryncha and is seen as a monophyletic group. In addition, Sternorryncha has been named Euhemiptera because was supported as a sister group to all the other hemipterans. Regarding Auchenorryncha, this lineage has been divided into Fulgoromorpha and Cicadomorpha, and controversy of whether is monophyletic or not has been argued. Concerning Heteroptera, other hemipteran phylogeny views argue about which group, the Fulgoromorpha, Cicadomorpha or Coleorrhyncha, is more closely related to this group. However, studies have shown the monophyly of Fulgoromorpha, Cicadomorpha and Heteroptera, and molecular evidence of both 18S DNA and complete mitogenomes support the paraphyly of Auchenorrhyncha (Cui et al. 2013). Furthermore, recent cladistics analysis of morphological characters and partial 18S rDNA sequence data determined that Homoptera is not a monophyletic group (*Cui* et al. 2013; Song et al. 2012). Nevertheless, despite all the different hypothesis, the phylogenetic relationships of Hemiptera remain unclear (Cui et al. 2013). Still, the monophyly of the order is strongly supported by the highly specialized sucking mouthparts of its members, and to reinforce the monophyly of the order, as a second character, it was proposed that the fork of the anterior axillary fold-line of the forewing should be considered as another synapomorphy (Forero 2008; Yoshizawa and Saigusa 2001).



1	Monophyletic Euhemiptera:	Hennig 1969, 1981; Kristensen, 1975, 1991; Schuh, 1979; Popov, 1981; Wootton & Betts, 1986; Minet & Bourgoin, 1986; Zrzavý, 1990, 1992; Wheeler *et al*., 1993; Campbell *et al*., 1994, 1995; von Dohlen & Moran, 1995; Sorensen *et al*., 1995; Ouvrard *et al*., 2000; Cryan & Urban, 2011
	Monophyletic Homoptera:	Macleay,1821; Schiödte, 1870; Handlirsch, 1908; Comstock, 1925; Brues & Melander, 1932; Song *et al*., 2010
2	Monophyletic Auchenorrhyncha:	Szelegiewicz, 1971; Kristensen, 1975, 1991; Hennig, 1981; Carver *et al*., 1991; Yoshizawa & Saigusa, 2001; D'Urso, 2002; Cryan & Urban, 2011
	Paraphyletic Auchenorrhyncha:	Goodchild, 1966; Hamilton, 1981; Bourgoin, 1988, 1993; Campbell *et al*., 1994, 1995; Sorensen *et al*., 1995; von Dohlen & Moran, 1995; Ouvrard *et al*., 2000; Xie *et al*., 2008
3	Monophyletic Heteropterodea:	Schlee, 1969; Zrzavý, 1990, 1992; D'Urso 1993; Campbell *et al*., 1995; Sorensen *et al*., 1995; Ouvrard *et al*., 2000; Grimaldi & Engel, 2005; Xie *et al*., 2008; Cryan & Urban, 2011
	Paraphyletic Heteropterodea:	Cobben, 1978; Shcherbakov, 1988; Bourgoin,1988, 1993; Popov & Shcherbakov, 1991, 1996

Figure 1.1. A summary of the different phylogenetic relationships proposed for the higher-level hemipteran lineages (Cui et al. 2013).

Psyllids (jumpling plant lice)

The psyllids (jumpling plant lice – Figure 1.2) belonging to the superfamily Psylloidea are a group of approximately 3800 species that are small plant-sap-feeding insects (Burckhardt 2014; Lubanga et al 2014; Hodkinson 2009), and host-specific, especially as larvae (Drohojowska and Burckhardt 2014). Their geographical range includes all the world's major zoogeographical regions (Drohojowska and Burckhardt 2014; Lubanga et al. 2014; Hodkinson 2009). The earliest psyllid ancestors are known from the Permian, and their diversification seems to be related to the angiosperms diversification in the Cretaceous. Moreover, the closest psyllid species to present day form were present in the Miocene and it is thought that many of the major known contemporary species of the group had a southern origin, probably on Gondwana during the Cretaceous (Hodkinson 1989).

Figure 1.2. Lateral view of an adult psyllid (jumping plant lice) showing the extended hind legs adapted for leaping (Barclay 2015).

Psyllids Biology

Psyllid reproductive biology and mating

Males and females of some psyllid species reach reproductive maturity after 24 to 48 hours of eclosion (Lubanga et al. 2014). Psyllids are frequently found highly aggregated or in mixed-sex colonies on the tissues of their host (Hodkinson 2009), although females are often solitary or found in loose aggregations. When finding a mate, in most species, competition between males takes place but without physical combat (Lubanga et al. 2014). In addition, during courtship, some species are known to produce drumming sounds or vibrations as they rub their forewings (Lubanga 2014; Hodkinson 2009); and studies of vibrational communication have shown that for some species there is a single copulation and for some others, females need to mate multiple times. Likewise, this same type of studies have found that females respond only when they perceive a male signal (Lubanga et al. 2014).

Recently, semiochemical signaling, long-range chemical signals for communication purposes, was described to play a role in psyllid mate attraction. Four psyllid species are known to be attracted to female-produced semiochemicals, including: *Cacopsylla pyricola, C. bidens, Bactericera cockerelli* and *Diaphorina citri*. The use of semiochemicals differs depending on the psyllid species. For instance, there are some cases of repulsion or attraction between the same genders. Pheromone glands have not been identified in psyllids yet, but cuticular hydrocarbons (complex of n-alkanes, methyl-branched alkanes and alkalenes with chains of approximately 21

to 37 carbon atoms) likely aid in mate attraction when biologically active. These chemicals then are perceived, at short range, by chemoreception (Lubanga *et al*. 2014). After mating, oviposition will usually take place within 24 hours (Lubanga et al. 2014). On average, female psyllids can laid from 40-50 to over 1000 eggs, but the majority of species lay between 200-800 eggs. The highest values are found often in pest species of Psyllidae and Triozidae (*Agnoscena pistaciae*, *Trioza erytreae* and *Bactericera cockerelli*) in which one female can produce up to 1300 eggs (Hodkinson 2009).

Life cycle determinant factors

Mainly, psyllids are associated with perennial dicotyledenous angiosperms, although there are a few species that develop on monocots and on gymnosperms. Also, adults from most of the psyllids species exploit conifers as shelter plants for overwintering times (Hodkinson 2009). Since the number of plants on which a psyllid completes its life cycle is usually one or a few closely related plant species, usually within a plant genus or family (Hodkinson 1989), this insect often become pests of cultivated plants. Psyllids transmit several bacterial species such as Liberibacter and Phytoplasma, causal agents of economic important plant diseases on citrus, pear, apple, plum, potato and tomato (Drohojowska and Burckhardt 2014).

The life cycle of psyllids involves an egg stage, five larval instars and a sexually reproducing adult stage, showing a 1:1 sex ratio at emergence (Lubanga et al. 2014; Hodkinson 2009). The nymphal biologies include species with nymphs that do not develop beneath a formed shelter or in a gall, called free-living; nymphs that induce and develop inside a plant gall, the gall-forming nymphs; nymphs that develop beneath a shelter of their own making, the lerp-forming nymphs; and in some cases, the inquiline nymphs, those that reside beneath shelters or in galls made by other species (Lubanga et al. 2014). Psyllid life cycle completion is strongly determined by temperature and precipitation. These two factors are a major determinant of egg and larval development, and so too is the potential number of generations per year of a species, also known

as the voltinism. In warm and wet conditions, psyllids life cycles are continuous with multiple generations. Nevertheless, psyllids have modified their physiognomy, physiology and phenology, shaping their life cycle to adapt to a wide range of climatic zones and exploit an extensive range of host plants.

Tropical and subtropical species (*Heteropsylla cubana, Diaphorina citri, Trioza erytreae* and *Trioza magnicauda*) with several generations in a season can complete between 8 and 16 generations per year; others at high elevations, like *Strophingia ericae,* take 2 years to complete one generation. On the other hand, temperate species are univoltine (having one breed in a season) and their development period is synchronized with the growing season of their host plant. Similarly, arctic genera such as *Cacopsylla, Psylla* and *Bactericera*, despite their adaptations to cold climates, show a similar trend to univoltism. In addition, some of the widely distributed species (e.g. *Trioza cinnamomi*) exhibit only one generation in the cooler parts of their range but can have multiple generations in other regions. However, life cycle completions of psyllid species of tropical, subtropical and dessert ecosystems can be affected when psyllids are exposed to elevated temperatures and reduced humidity. Both factors have a major influence, affecting fecundity, mortality, development, and range limits (Hodkinson 2009).

Another adaption is seen in species of psyllids living in seasonal environments, were they have successful life cycle completions. Success is predicated on their ability to synchronize their development with that of their host plant by slowing down their growth, dependant on environmental factors such as photoperiod and temperature. This cessation or slowing down of development, called diapause, provides the right timing of development and can happen at any of the life stages: egg, larval or adult stages. Link to diapause are metabolism adaptations, in which basal metabolism is minimize during seasonal inactivity. Likewise, since the metabolic energy cost is influenced by the environmental temperature, body size and sex; adult male psyllids tend to be smaller than females because they have a higher metabolic rate (Hodkinson 2009).

One more important factor in psyllids life cycle completion is the host species. Breeding success can be different depending on the potential host species or among varieties of the same host species. For instance, different plant species in a given host range can be either highly susceptible to psyllid species or resistant to them. Similarly, these insects can occur within varieties and cultivars of the same host species. This variation on host species is explained by differences in how attractive the foliage is, and also the oviposition and larval survival rates, as well as the larval developmental period. Although for some psyllid species, the host plant that is most attractive will not necessarily be the appropriate plant for larval development. In addition, other factors that help to determine a preference to a host plant are: the host species phenology, leaf color, the glaucous wax on the leaf surface, the physical hardiness of the terminal shoot, the presence of attractive chemicals like caryophyllene, and low concentrations of repellent chemicals (phenolics, terpenoids or glucosinates). Despite the fact that the host plant provides all the elements for the psyllids growth, when mature, leaves no longer provide a good source of soluble nutrients, especially nitrogen in the form of amino acids. However, psyllids can enhance the availability of nutrients for the larval growth and development, by injecting saliva into the phloem or into the leaf mesophyll. The amylase then, induces localized symptoms, resulting in cell degeneration that resembles to premature senescene, which promotes the mobilization of lipids, amino acids and soluble proteins in the area. This enhancement can be increased when psyllids feed in groups, but depending on the species, high densities can lead to a decline of the host plant quality. Psyllid species known as important agricultural pest (*Trioza apicalis* on carrot and *Bactericera cockerelli* on potato, tomato and pepper), cause damage in their host plant, because instead of improving the availability of the host plant's nutrients, they incite a systemic phytotoxaemia (Hodkinson 2009).

Dispersal

Psyllids are effective dispersers in both short and long distances, but are also assisted by the wind. This dispersal ability is a key element in their life history because it allows to track the different spatial distribution of their host plants, as well as to move between different host plant species and to exploit others for overwintering. Psyllids are so effective as dispersers that some genera have been found in drogue nets behind aircrafts or in kite mounted nets; while others are common components of aerial deposition. Moreover, being able to disperse, is favorable for psyllids because it counteracts intraspecific competition and natural enemies (Hodkinson 2009).

In psyllids, the number of generations per year, influences the dispersal behavior. Multivoltine species can show short dispersal or long dispersal behavior depending on the season; and for some species it gives the opportunity for host plant alternation. Only two known multivoltine species are known for doing this: *Bactericera crithmi* and *Bactericera cockerelli*. Both species overwinter in a different plant than its normal host. On the other hand, univoltine species show an autumn and spring peak dispersal as they move to and from their overwintering hosts (Hodkinson 2009).

Overwintering

Psyllid species overwinter when its host is dormant or it is an unfavorable period for the insect development. In general, temperate adult psyllids overwinter on evergreen shelter plants and move back to their true host on spring. This evergreen shelter plants include conifers from the species of *Pinus*, *Picea*, *Abies*, *Taxus*, *Tsuga*, *Cupressus*, and *Juniperus* and sometimes also *Ulex*. However, it is not known if psyllids feed on the shelter plants. Furthermore, some temperate psyllid species experience an adaptive strategy, ovarian diapause, while they are overwintering. Genera such as *Aphalara*, *Livia*, *Cacopsylla*, *Bactericera*, *Phylloplecta* and *Trioza* show this strategy (Hodkinson 2009).

Psyllids as vectors of plant pathogens

Insects with piercing-sucking mouthparts, like psyllids, are major vectors of plant pathogens and are responsible of transmitting more than 50% of vector-borne viruses and bacterial pathogens in plants (Huot et al. 2013). Some psyllids species are well known because of their capacity to transmit pathogens to their host plants; especially when nymphs feed, since they inject toxic secretions through saliva, resulting in serious symptoms to the host plants. Usually, psyllids show close association with these pathogens, because this association increase the chance of inducing changes in the host plant, such as physiological changes or reduce levels of defensive chemicals. Therefore, benefiting the psyllids nutrition (Hodkinson 2009).

Psyllids can transmit viruses and bacteria (liberibacters, phytoplasmas and fireblights – *Erwinia amylovora*-), which are ingested by the insect during feeding and later reinjected into the host plant within the psyllid saliva (Hodkinson 2009). Both larvae and adult are capable of transmiting phytoplasmas and evidence exists that this bacteria can be transferred from female psyllids to their offsprings, e.g. *Cacopsylla pruni*, (Hodkinson 2009). In the case of liberibacters, they are known to be transmitted by at least four psyllid species (*Diaphorina citri, Trioza erytreae, Bactericera cockerelli* Sulc and *Trioza apicalis*) and is associated with economically and newly important diseases affecting citrus and solanaceous crops (Munyaneza 2010). However, unlike phytoplasmas, the liberibacter does not seem to be transmitted vertically (from mother to offspring) between generations (Hodkinson 2009).

Endosymbionts in psyllids species

Besides the pathogens, viruses and bacteria, psyllids have associations with endosymbiotic bacteria which plays a direct role in psyllids nutrition by synthesizing the essential amino acids and vitamins from the sugar rich phloem sap (Hodkinson 2009; Thao et al. 2000). Also, mutualistic bacteria can enhance insects' resistance to pesticides, enhance their immune response and give reproductive advantage (Arp et al. 2014). This bacteria is found within a

specialized cell called bacteriocyte that groups together to form a bacteriome within the insect's body cavity (Hodkinson 2009; Thao et al. 2000).

Psyllids present two main groups of endosymbionts: the primary (P), found in the bacteriocytes and because they are genetically similar throughout the psyllids, they likely have colonized just once and co-evolve with their host. *Candidatus* Carsonella ruddii is the P endosymbiont of psyllids. The other group is known as the secondary (S) endosymbiont, present in cells associated to the bacteriocytes, but its function has not been understood yet. In contrast with the P endosymbionts, these S endosymbionts include several different group within the Eubacteriaceae and some are associated more with parasitism than with nutrition (Hodkinson 2009). All of these endosymbionts are maternally-transmitted to progeny by cospeciation between the host and the endosymbiont (Thao et al. 2000).

The Potato Psyllid

The potato psyllid, *Bactericera* (=*Paratrioza*) *cockerelli* (Sulc), or the tomato psyllid as is also known, is a hemipteran insect belonging to the Triozidae family. Originally, this insect was described as *Trioza cockerelli* in 1909 by Sulc, but in 1911 was assigned to *Paratrioza*. Later, in 1997, *Paratrioza* was synonymized with the genus *Bactericera* and the potato psyllid changed not only of genus but also of family, from Psyllidae to Triozidae (Butler and Trumble 2012).

The potato psyllid has a wide natural host range and is considered a serious pest of potato and other solanaceous crops (Buchman et al. 2012; Butler and Trumble 2012; Nachappa et al. 2012; Munyaneza et al. 2007), but can also survived on plants of the family Convolvulaceae, which support the normal development of the psyllids (Rondon et al. 2012). The potato psyllid is native to the southwestern United States and northern Mexico (Buchman et al. 2012; Nachappa et al. 2012); but recently, has been documented that its geographical range has increased, including

the northwestern and western regions of the US, some countries in Central America (as far south as Honduras) and New Zealand (Butler and Trumble 2012; Nachappa et al. 2012).

The different outbreaks of the pest, associated with the absence of elevated temperatures, have allowed the recognition of different potato psyllid biotypes and haplotypes (Swisher et al. 2014; Chapman et al. 2012; Munyaneza 2012; Swisher et al. 2012) by using mitochondrial genes such as cytochrome oxidase I (COI) as a genetic marker (Swisher et al. 2013; Chapman et al. 2012; Munyaneza 2012; Swisher et al. 2012; Liu et al. 2006). The biotypes identified as western (California, New Mexico up to Washington, Oregon and Idaho) and central (eastern Mexico up to Texas, Kansas, Colorado, Nebraska, and Wyoming) differ genetically, but also are biologically different, in terms of survivorship, growth index and developmental time (Swisher et al. 2013; Munyaneza 2012; Swisher et al. 2012; Liu et al. 2006). Whereas, the northwestern (Washington, Oregon and Idaho) and southwestern (predominantly in New Mexico and Southern Colorado) are considered haplotypes because, so far, only their genetic differences have been identified by using a single mitochondrial gene (Swisher et al. 2013).

Biology of the Potato Psyllid

Potato psyllids are polyphagous insects with a wide array of hosts, more than 20 plant families, and more than 40 host species where it can oviposit and complete its life cycle. The preference of the host is associated with abundance, preference and to those in close proximity to agricultural areas, but in general have a preference from species from the Solanaceae family (Butler and Trumble 2012). There is evidence that adult psyllids migrations are how this species arrive in agricultural crops (Butler and Trumble 2012). Adult psyllids have been collected via airplane at altitudes up to 1524m; and have been captured together with the sugar beet leafhoppers, which is known to migrate from southern areas. Also, *B. cockerelli* intolerance to the high temperatures of their breeding sites during summer and to the winter temperatures of north and central states is considerably enough to believe that they migrate from one place to another

(Liu et al. 2006). Potato psyllids overwinter from California to south of Texas and north of Mexico (Rondon et al. 2012), as well as in Washington, Oregon and Idaho where it has been found overwintering on the weedy *Solanum dulcamara* (Swisher et al. 2013). In addition, this species can overwinter on the stage of egg, larva or as adults and overwinters in the leaves of the host plants, either on conifers or evergreen shrubs (Hodkinson 2009).

The biology of the potato psyllid is influenced by the temperature and humidity. *B. cockerelli* preferred to live in temperate dry zones (Hodkinson 2009) and their life history characteristics are impacted by extremely hot or cold conditions (Butler and Trumble 2012). Nevertheless, under optimal conditions, psyllids can complete a generation in less than a month, with usually 3 or more generations per year (Hodkinson 2009). *B. cockerelli* have three life stages: egg, nymph and adult. One female can deposit about 300 to 500 eggs (Abdulla 2013; Wallis 1955). The eggs (Figure 1.3) are yellow-orange with a football type shape and extremely small. Eggs are usually laid in the underside and along the edges of leaves, and each one of them attaches individually either to the leaf or to a short stalk. The hatching of the eggs occurred between 6 to 10 days, depending on the temperature. Warmer temperatures favor early hatching, although temperatures above 32°C reduce reproduction and survival (Rondon et al. 2012).

The nymphs (Figure 1.3) are flat (Rondon et al. 2012) and orange to yellow when newly hatched, but become green pale when mature (Wallis 1955). The nymphs present a fringe of short spines or hairs around the edge of the body. However, larger nymphs have distinct wingpads on their dorsum. In warm temperatures, immature psyllids go through five stages or instars in as few as 13 days. Nymphs are usually found in the underside of leaves and look like whiteflies or scales, but differs from them in that they move readily when are disturbed (Rondon et al. 2012).

Figure 1.3. Eggs and nymphs of B. cockerelli Sulc (Rondon et al. 2012).

The adults (Figure 1.4) are approximately 2 mm long and have clear wings that rest roof-like over the body, resembling as small cicadas, winged aphids, or bark lice (Rondon et al. 2012). When newly emerged are light green (Wallis 1955), making it difficult to distinguish from other insects, but after two or three days adults are predominantly black with white markings (Rondon et al. 2012), which gives them a gray appearance (Wallis 1955). The first abdominal segment present a broad white band and the last segment has an inverted white V. Adults are very active, jumping quickly when disturbed (Rondon et al. 2012).

Figure 1.4. Potato psyllids (B. cockerelli Sulc) adults showing the whitish bands on the abdomen and the white or yellow lines on the head and thorax (Munyaneza 2012).

Potato psyllid endosymbionts

Potato psyllids as plant phloem sap feeders lack nitrogenous compounds in their diet. Therefore, the presence of prokaryotic intracellular symbionts or endosymbionts in potato

psyllids gut, supplement those excluded essential nutrients in their diet, for exchange of permanent residence in specialized cells called bacteriomes (Arp et al. 2014; Hail et al. 2012; Thao et al. 2000). Insects can have a primary and a secondary endosymbiont (Thao et al. 2000). For instance, *Candidatus* Carsonella ruddii is the primary endosymbiont of the potato psyllid. This bacteria has likely coevolved with psyllids and provides those nutrients that the insect is lacking (Arp et al. 2014; Hail et al. 2012; Thao et al. 2000). In contrast, the secondary endosymbionts can be acquired by the insect through feeding or passed to offspring transovarially, examples are *Candidatus* Liberibacter solanacearum, *Acinetobacter*, *Methylibium*, and *Wolbachia* sp. Although the purpose of having secondary endosymbionts is unknown, some can provide physiological alterations. Perhaps, give resistance to insecticides (Arp et al. 2014; Hail et al. 2012). Microbial community studies of the potato psyllid have indicated that these endosymbionts are influenced primarily by location, but not by host plant or haplotype (Arp et al. 2014).

Potato psyllid and plant diseases (direct and indirect damage)

The potato psyllid is a polyphagous phloem feeder (Alvarado et al. 2012; Nachappa et al. 2012; Ramírez-Davila et al. 2012), and phloem-feeding insects require symbiotic relationships with bacteria to supplement missing nutrients, like amino acids, that are not found in the plant sap (Alvarado et al. 2012; Hail et al. 2012). A phloem feeder, such like the potato psyllid, can cause two types of damage, a direct damage caused by the nymphs and a second one considered as indirect damage. While potato psyllids feed, the nymphs have been reported to inject a toxin, which causes plants' yellowishing and underdevelopment known as the "psyllid yellows disease", which consequently affects tuber yield and quality (Munyaneza et al. 2007; Ramírez-Davila et al. 2012; Liefting et al. 2009). Although, these symptoms can also result from tissue damage or from the lack of certain nutrients in the plant. The indirect damaged, occurs with the transmission of a phytoplasma by adults and nymphs, which is the causal agent of the purple top potato disease.

Recently, the potato psyllid has been associated with the Zebra Chip disease, transmitting the causal agent *Candidatus* Liberibacter solanacearum (syn. Ca. L psyllaurous) (Ramírez-Davila et al. 2012; Buchman et al. 2011; Liefting et al. 2009). Once the insect, adult or nymph, acquires the bacterium from feeding on an infected plant; the insect always carries the bacterium. Besides, a percentage of the young hatching from eggs lay by an infected adult become carriers of the bacterium as well (Rondon et al. 2012).

Karel Sulc, in 1909, predicted that the potato psyllid may become a destructive pest due to the large amount of nymphs of potato psyllids he observed in peppers in Boulder, Colorado. Six years later, this species was recognized as a plant pest when it damaged for the first time the False Jerusalem Cherry, *Solanum capsicastrum,* in some areas of California Then in 1927, outbreaks of psyllid yellows occurred on potatoes in different states of USA, and then again in 1938 (Butler and Trumble 2012). Nevertheless, historically, potato psyllid infestations were sporadic (Butler and Trumble 2013; Hail et al. 2012), causing damages for six months or a year and then disappearing for 20 or 30 years (Butler and Trumble 2013). However, in recent years the outbreaks have become much more persistent with year-long and multiyear infestations becoming the norm (Hail et al. 2012). In 1994, a new potato disease was discovered in Mexico and later named as "zebra chip" (Butler and Trumble 2012). Then, during 2001 a series of outbreaks, recurring each year, were noted in tomato and potato crops in the western and northwestern US, western Mexico and in Canada. In 2006, the psyllid became a pest in potatoes in Guatemala and Honduras (Abdullah 2008); spreading, also, to New Zealand (Butler and Trumble 2013). Also they became a limiting factor in the production of potatoes in both Central and North America (Butler and Trumble 2013; Hail et al. 2012). Surveys of insects associated with the potatoes crops in affected areas of the southwestern United States indicated that the psyllid was the most common and abundant insect in all Zebra chip infected potato fields (Munyaneza et al. 2007). The observed expansion in the geographical range of the potato psyllid coincided with the discovery of its transmission ability of a bacterial plant pathogen (Nachappa et al. 2012).

Zebra Chip Disease

Zebra Chip is a disease that was originally observed in potato crops in Mexico and the United States: Texas, California, Colorado, Kansas, Nebraska and New Mexico (Munyaneza et al. 2007; Lin et al. 2009; Secor et al. 2009), and recently in New Zealand (Liefting et al. 2009) and most of Central America (Secor and Rivera-Varas 2004; Munyaneza et al. 2007; Secor et al. 2009). This disease was first identified in 1994 in Texas and Mexico, where it is known in Spanish as "papa manchada" or potato stained (Crosslin and Munyaneza 2009; Secor et al. 2009).

The impact and severity of the disease varies year to year and it is accompanied by serious economic consequences because it reduces the commercialization of the fresh potatoes for chips (Lin *et al*. 2009), causing storage losses and limiting exports (Guenthner et al. 2012); as well as the abandonment of whole areas dedicated for potato growth (Gao et al. 2009). The disease is named Zebra Chip because the tubers present very superficial stripes on different directions (Calderoni 1978). Zebra Chip is identify by a pattern of stripes or necrotic points that are formed in the tuber, especially when they are processed to produce potato chips, converting starch in sugars which turned dark when are cooked (Munyaneza et al. 2007).

The infected plants show similar symptoms to the purple top disease and psyllid yellows disease, so specific symptoms differentiating Zebra Chip disease must be used. The symptoms include: chlorosis, twisted stems with a zig-zag appearance, enlarges nodes, aerial tubers, vascular discoloration and, leaves that arrive prematurely to senescence (Crosslin and Munyaneza 2009; Lin et al. 2009; Secor et al. 2009). Also, the tubers of infected plants present sunken stolons, frequently with a pink color (Crosslin and Munyaneza 2009). However, the primary symptom and the one that separates the disease from others, is a brown coloration in the vascular ring and medular rays, which extends throughout the tuber. Additionally, the tubers show poor or null germination as a result of lenticels death (Lin et al. 2009). The disease prompts tuber necrosis and compromises the phloem vessels (Secor et al. 2009) and constitutes a lethal condition for the potato plants (Wen et al. 2009).

Zebra Chip is a complex disease that involves the potato psyllid, *Bactericera cockerelli* Sulc (Hansen et al. 2008; Liefting et al. 2009), as the vector of the Zebra Chip pathogen, and results of psyllids infestations correlate to the development of the disease. One of the primary effects of this disease is the reduction of photosynthesis that alters the sugar metabolism within the tuber (Gao et al. 2009). The putative causal agent of the disease is a new species of *Candidatus* Liberibacter: *Ca.* L. solanacearum (Lso), also known as *Ca*. Liberibacter psyllaurous (Munyaneza 2012; Hansen et al. 2008; Liefting et al. 2009); but its role in the expression of the disease is not confirmed (Crosslin and Munyaneza 2009), because the pathogen is not culturable so traditional Koch's postulates cannot be completed. However, in the last years Lso has also been detected on carrots from Finland, Norway, Sweden and the Mediterranean Region, as well as in celery crops from Spain; implying that the bacteria has more than one host and more than one vector (Munyaneza 2012).

Candidatus Liberibacter solanacearum (Lso)

In 2008, *Candidatus* Liberibacter solanacearum (Lso) was first described in New Zealand, and later was found in the USA (Liefting et al. 2009; Janse 2012; Munyaneza 2012). The identification of the causal agent of the Zebra Chip disease began with the emergence of a disease like phytoplasma in glasshouse-house grown tomato and pepper from New Zealand. Transmission electron microscopy on symptomatic tomato leaf tissue revealed the presence of a phloem-limited bacterium like organisms (BLOs). PCR assays, sequencing and phylogenetic analysis of the 16S region defined the new species of bacterium as a liberibacter, and was considered as a new species of the genus "*Candidatus* Liberibacter". The presence of this new BLOs species in potatoes was supported with the detection of the bacterium in infected tuber, thus associating the new species with the zebra chip disease and naming it "*Candidatus* Liberibacter solanacearum". At the same time, in the US, a potato psyllid symbionts study was taken place, and as a result, a bacterium was designated as "*Candidatus* liberibacter psyllaurous"

because of its relationship to psyllid yellows disease. Further research involving the redesigning of PCR primers based on a related *Candidatus* liberibacter species, *Ca.* L. asiaticus, confirmed the presence of "Los" in potato psyllids reared on symptomatic plants and from the wild. Later, sequence analysis evidenced that *Ca*. L. psyllaurous and solanacearum are the same bacterium (Crosslin et al. 2010).

Ca. liberibacter solanacearum is closely related to the liberibacters that are associated with the citrus greening disease. The liberibacters are bacteria that belong to the *Alphaproteobacteria* and are characteristic for being limited to phloem, gram-negative and are unculturable (Munyaneza 2012; Hansen et al. 2008; Liefting et al. 2009). Lso is rod-shaped, about 0.2 µm wide and 4 µm long (Munyaneza 2012) and adapts to different climates such as desert, steppe, Mediterranean, marine coast, humid continental and humid subtropical (Janse 2012). Hence, this pathogen's geographic distribution includes western and central region of USA, Mexico, Central America and New Zealand, and has been documented in northern Europe and the Mediterranean region (Munyaneza 2012). Lso has also a wide array of hosts, including: pepper (*Capsicum annuum*), chili pepper (*C. frutescens*), tomato (*Lycopersicon esculentum),* tomatillo (*Physalis peruviana*), tamarillo (*Solanum betaceum*), potato (*S.tuberosum*), tobacco (*Nicotiana tabacum*), eggplant (*S. melongena*) and several weeds in the Solanaceae family (EPPO 2013; Janse 2012; Munyaneza 2012). Besides, it has also been documented on carrots from Finland, Norway, Sweden, and the Mediterranean region; as well as on celery in Spain (Munyaneza 2012). Lso is spread from infected to healthy plants (Munyaneza 2012), through potato seed, tomato plantlets, fruits, grafting and by psyllid vectors (Janse 2012). Depending on the region, the bacteria has different psyllid vectors such as *Trioza apicalis* (Finland), *Bactericera trigonica* (Canary Islands and Spain), and *Bactericera cockerelli* (USA, Central America and New Zealand) (Janse 2012; Munyaneza 2012).

Analysis of single nucleotide polymorphisms on the 16S rRna, 16 s/23 s ISR and 50s rplJ and rplL ribosomal proteins genes of the bacteria, identified four haplotypes: A, B, C and D,

by their geographic distribution. Haplotype A is known from Honduras, Guatemala, western Mexico to Arizona and California, and in New Zealand. Haplotype B has been found from eastern to the north of Mexico, Texas and south central Washington. These two haplotypes, A and B, are associated to diseases caused in potatoes and other plants of the Solanaceae family. Haplotypes C and D, are associated with the disease on infected carrots. Haplotype C has been found in northern Europe (Finland, Sweden and Norway), and haplotype D in Spain and Canary Islands (EPPO 2013; Munyaneza 2012). Although, Lso adapts to different climate zones, it appears to be heat-sensitive, been its optimal developmental temperature of zebra chip disease symptoms on potato plants of 27-32°C. Also, within the adult potato psyllid, the bacteria is sensitive to high and low temperatures, preferring a temperature similar to that of the potato psyllid temperature reproduction. In addition, Lso is equally found in potato psyllid males and females. Nonetheless, concentration is less in fifth-instar nymphs, but increases in the adults (Munyaneza 2012).

The economic importance of the potato crop

The potato, *Solanum tuberosum*, is the fourth most important crop in the world after maize, rice and wheat (Thiele et al. 2008; Askew 2001). It is consider a delicacy, a fast food and a hedge against famine, because potatoes are not only nutritious, but also relatively low in calories, virtually free of fat and cholesterol, and high in vitamin C and potassium. Also, are high in fiber when serve with their skins (Mcgregor 2007).

The potato has been used since 7 to 10,000 years ago in the Lake Titicaca region (Peru/Bolivia), and later with the arrival of Europeans to America, it became a staple crop in Europe and the United States. The crop was introduced to Asia by early 17th century and, since the second half of the 20th century it has beginning to play a greater role in the developing countries (Thiele et al. 2008). The reason of potato crop success is its versatility. The crop grows in different climatological zones: temperate regions, sub-tropics and tropics; and also under vary

different agro-ecological conditions, lowlands and highlands and in very different socio-economic environments across the world (Askew 2001).

All over the world, the potato crop is of vital importance to the small farmers and their families since it represents a good source of nourishment and income. Unlike major cereals, the production of potato is one of the most interesting: it yields between two or four times more than rice or wheat and 85% of the plant is edible (Devaux and Ordinola 2012). Also, it is not a globally traded commodity and the potato prices are determined by local production costs. Therefore, is a highly recommended food security crop that can help farmers with low income and vulnerable consumers to overcome the crisis due to the world supply and demand (FAO 2008). However, diseases in potato can generate high economic losses in its production, which can result in a limiting factor, especially in all of those countries that depend economically on agriculture.

References

Abdulla, N.M.M. 2008. Life history of the Potato Psyllid *Bactericera* cockerelli (Homoptera: Psyllidae) in controlled Environment agriculture in Arizona. African Journal of Agricultural Research 3(1): 060-067.

Alvarado V.Y., D. Odokonyero, O. Duncan, T. E. Mirkov, H.B. Scholthof. 2012. Molecular and Physiological Properties Associated with Zebra Complex Disease in Potatoes and Its Relation with Candidatus Liberibacter Contents in Psyllid Vectors. PLoS ONE 7(5): e37345

Arp, A., J. E. Munyaneza, J. M. Crosslin, J. Trumble, and B. Bextine. 2014. A Global Comparison of *Bactericera cockerelli* (Hemiptera: Triozidae) Microbial Communities. Environ. Entomol. 43(2): 344-353.

Askew, M. F. 2001. The economic importance of the potato. Virus and virus-like diseases of potatoes and production of seed-potatoes. 1st Ed. G. Loebenstein *et al*. the Netherlands, Kluwer Academic Publishers, 1-18 pp.

Barclay, C. 2015. Infested plants – development and growth. Barclay Gardening. Retrieve from: http://www.barclaygardening.us/infested-plants/development-and-growth.html

Buchman, J., T. W. Fischer, V. G. Sengoda and J. E. Munyaneza. 2012. Zebra Chip Progression: From Inoculation of Potato Plants with Liberibacter to Development of Disease Symptoms in Tubers. Am. J. Pot Res 89:159-168.

Buchman, J. L., V. G. Sengoda and J. E. Munyaneza. 2011. Vector Transmission Efficiency of Liberibacter by *Bactericera cockerelli* (Hemiptera: Triozidae) in Zebra Chip Potato Disease: Effects of Psyllid Life Stage and Inoculation Access Period. J. Econ. Entomol. 104(5): 1486-1495.

Burckhardt, D., D. Ouvrard, D. Queiroz and D. Percy. 20014. Psyllid host-plants (Hemiptera: Psylloidea): Resolving a semantic problem. Florida Entomologist 97(1): 242-246.

Butler, C. and J. Trumble. 2013. Potato Psyllid. Center for Invasive Species Research (CISR) of the University of California, Riverside. Retrieved from: http://cisr.ucr.edu/potato_psyllid.html

Butler, C., and J. T. Trumble. 2012. The potato psyllid, *Bactericera cockerelli* (Sulc) (Hemiptera: Triozidae): life history, relationship to plant diseases, and management strategies. Terrestrial Arthropod Reviews, 5: 87-111.

Chapman, R. I., J. F. Macias-Velasco, A. P. Arp, and Blake R. Bextine. 2012. Using quantitative real time PCR melt curve analysis of partial CO1 sequence for rapid biotype differentiation of *Bactericera cockerelli* (Hemiptera: Triozidae). Southwestern entomologist, 37(4): 2012.

Crosslin, J. M., J. E. Munyaneza, J. K. Brown, L. W. Liefting. 2010. Potato zebra chip disease: a phytopathological tale. Plant Health Progress. Retrieved from: http://www.plantmanagementnetwork.org/pub/php/review/2010/zebra/

Crosslin, J. M. and J. Munyaneza. 2009. Evidence that the Zebra Chip disease and the putative causal agent can be maintained in potatoes by grafting and In Vitro. American Journal of Potato Research (AM. J. Pot Res), 86(3):183-187.

Cryan, J. R. and J. M. Urban. 2012. Higher-level phylogeny of the insect order Hemiptera: is Auchenorrhyncha really paraphyletic?. Systematic Entomology 37: 7-21.

Cui, Y., Q. Xie, J. Hua, K. Dang, J. Zhou, Z. Liu, G. Wang, X. Yu, and W. Bu. 2013. Phylogenomics of Hemiptera (Insecta: Paraneoptera) based on mitochondrial genomes. Systematic Entomology, 38: 233-245.

Devaux, A. and M. Ordinola. 2012. De América para el mundo. Revista Papa. 16-19 pp.

Drohojowska, J., and D. Burckhardt. 2014. The jumping plant-lice (Hempiter: Psylloidea) or Turkey: a checklist and new records. Turk J Zool, 38:559-568.

European and Mediterranean Plant Protection Organization (EPPO). 2013. "*Candidatus* Liberibacter solanacearum". Bulletin OEPP/EPPO, 43(2): 197-201.

Food and Agriculture Organization of the United Nations. 2008. The international year of the potato. Retrieved from: http://www.fao.org/potato-2008/en/aboutiyp/index.html

Forero, D. 2008. The systematics of the Hemiptera. Revista Colombiana de Entomología 34(1): 1-21.

Gao, F., J. Jifon, X. Yang, and T. Liu. 2009. Zebra chip disease incidence on potato is influenced by timing of potato psyllid infestation, but not by the host plants on which they were reared. Insect Science, 16:399-408.

Guenthner, J., J. Goolsby, and G. Greenway. 2012. Use and Cost of Insecticides to Control Potato Psyllids and Zebra Chip on Potatoes. Southwestern Entomologists, 37(3): 263-270.

Hail, D., S. E. Dowd, and B. Bextine. 2012. Identification and Location of Symbionts Associated with Potato Psyllid (*Bactericera cockerelli*) Lifestages. Environmental Entomology 41(1): 98-107.

Hansen, A. K., J. T. Trumble, R. Stouthamer, and T. D. Paine. 2008. A new Huanglongbing species, "*Candidatus* Liberibacter psyllaurous," found to infect tomato and potato, is vecterored by the psyllid *Bactericera cockerelli* (Sulc). AppL. Environ. Microbiol. 74(18): 5862-5865.

Hodkinson, I. D. 2009. Life cycle variation and adaptation in jumping plant lice (Insecta:Hemiptera: Psylloidea): a global synthesis. Journal of Natural History, 43(1-2):65-179.

Hodkinson, I. D. 1989. The biogeography of the Neotropical jumping plant-lice (Insecta: Homoptera: Psylloidea). Journal of Biogeography 16: 203-217.

Huot, O. B., P. Nachappa, and C. Tamborindeguy. 2013. The evolutionary strategies of plant defenses have a dynamic impact on the adaptations and interactions of vectors and pathogens. Insect Science, 20: 297-306.

Janse, J. D. 2012. Bacterial disease that may or do emerge, with (possible) Economic damage for Europe and the Mediterranean Basin: notes on epidemiology, risks, prevention and management on first occurrence. Journal of Plant Pathology, 94(4): S4.5- S4.29

Lee, W., J. Kang, C. Jung, K. Hoelmer, S. H. Lee, and S. L. 2009. Complete Mitochondrial Genome of Brown Marmorated Stink Bug *Halyomorpha halys* (Hemiptera: Pentatomidae), and Phylogenetic Relationships of Hemipteran Suborders. Mol. Cells, 28: 155-165.

Liefting, L. W., P. W. Sutherland, L. I. Ward, K. L Paice, B. S. Weir, and G. R. G. Clover. 2009. A new "Candidatus Liberibacter" species associated with diseases of Solanaceous crops. Plant Disease, 93 (3): 208-214.

Lin, H., H. Doddapaneni, J.E. Munyaneza, E. L. Civerolo, V. G. Sengoda, J. L. Buchman, and D. C. Stenger. 2009. Molecular characterization and phylogenetic analysis of 16S rRNA from a new "*Candidatus* Liberibacter" strain associated with Zebra Chip Disease of potato (*Solanum tuberosum* L.) and the potato psyllid (*Bactericera cockerelli* Sulc). Journal of Plant Pathology, 91 (1):215-219.

Liu, D., J.T. Trumble, and R. Stouthamer. 2006. Genetic differentiation between eastern populations and recent introductions of potato psyllid (*Bactericera* cockerelli) into western North America. Entomologia Experimentalis et Applicata 118: 177-183.

Lubanga, U. K., C. Guédot, D. M. Percy, and M. J. Steinbauer. 2014. Semiochemical and Vibrational Cues and Signals Mediating Mate Finding and Courtship in Psylloidea (Hemiptera): A synthesis. Insects, 5:577-595.

Nachappa, P., J. Levy, and C. Tamborindeguy. 2012. Transcriptome analyses of *Bactericera cockerelli* adults in response to "*Candidatus* Liberibacter solanacearum" infection. Mol Genet Genomics 287: 803-817.

Mcgregor, I. 2007. The fresh potato market. Potato biology and biotechnology: advances and perspectives. 1st Ed. D. Vreugdenhil. Oxford, UK. Elsevier, 3-26 pp. Print.

Munyaneza, J. E. 2012. Zebra Chip Disease of Potato: Biology, Epidemiology, and Management. Am. J. Pot Res 89: 329-350.

Munyaneza, J. 2010. Psyllids as vectors of Emerging Bacterial Diseases of Annual Crops. Southwestern Entomologists, 35(3): 471-477.

Munyaneza, J.E., J.M. Crosslin, and J. E. Upton. 2007. Association of *Bactericera* cockerelli (Homoptera: Psyllidae) with "Zebra Chip", a New Potato Disease in Southwestern United States and Mexico. Journal of Economic Entomology 100 (3): 656-663.

Ramírez-Davila, J.F., E. Porcayo-Camargo, J. R. Sánchez-Pale, and L.M. Vásquez-García. 2012. Spatial Behavior Comparison of *Bactericera cockerelli* Sulc. (Hemiptera: Triozidae) in Mexico. Neotro Entomol 41: 9-16.

Rondon, S. A. Schreiber, A. Jensen, P. Hamm, J. Munyaneza, P. Nolte, N. Olsen, E. Wenninger, D. Henne, C. Wohleb, and T. Waters. 2012. Potato Psyllid Vector of Zebra Chip Disease in the Pacific Northwest –Biology, Ecology, and Management. Pacific Northwest Extension Publication; Oregon State University, University of Idaho and Washington State University. 8 pp.

Secor, G. A. and V. V. Rivera-Varas. 2004. Emerging diseases of cultivated potato and their impact on Latin America. Rev. Latinoamericana de la papa (suplemento) 1:1-8.

Secor, G. A.; V. Rivera, J. Abad, I. Lee, G. Clover, L. Liefting, X. Li, and S. De Boer. 2009. Association of "*Candidatus* Liberibacter solanacearum" with Zebra Chip Disease of Potato established by graft and psyllid transmission, electron microscopy, and PCR. Plant Disease, 93(6):574- 583.

Song, N., A-P Liang, and C-P Bu. 2012. A molecular Phylogeny of Hemiptera Inferred from Mitochondrial Genome Sequences. PLoS ONE 7(11): e48778.doi:10.1371/journal.pone.0048778

Swisher, K. D., D. C. Henne, and J. M. Crosslin. 2014. Identification of a fourth haplotype of *Bactericera cockerelli* (Hemiptera: Triozidae) in the United States. J. insect Sci. 14(161): DOI: 10.1093/jisesa/ieu023

Swisher, K.D, J. E. Munyaneza, and J. M. Crosslin. 2013. Temporal and Spatial Analysis of Potato Psyllid Haplotypes in the United States. Environ. Entomol. 42(2): 381-393.

Swisher, K. D., J. E. Munyaneza, and J. M. Crosslin. 2012. High Resolution Melting Analysis of the Cytochrome Oxidase I Gene Identifies Three Haplotypes of the Potato Psyllid in the United States. Environmental Entomology, 41(4): 1019-1028.

Thao, M. L., N. A. Moran, P. Abbot, E. B. Brennan, D. H. Burckhardt, and P. Baumann. 2000. Cospeciation of Psyllids and their Primary Prokaryotic Endosymbionts. Applied and Environmental Microbiology, 66 (7): 2898-2905.

Thiele, G., G. Hareau, V. Suárez, E. Chujoy, M. Bonierbale, and L. Maldonado. 2008. Varietal change in potatoes in developing countries and the contribution of the International Potato Center: 1972-2007. International Potato Center (CIP), Lima, Peru. Working paper 2008-6. 6 pp.

von Dohlen, C. D., and N. A. Moran. 1995. Molecular Phylogeny of the Homoptera: A Praphyletic Taxon. J. Mol. Evol. 41: 211-223.

Wallis, R. L. 1955. Ecological Studies on the Potato Psyllid as a Pest of Potatoes. Technical Bulletin No. 1107, U. S. Department of Agriculture. Retrieved from: http://books.google.com/books?hl=en&lr=&id=s_fAFHpcG3oC&oi=fnd&pg=PA1&dq=biology+of+the+potato+psyllid&ots=liAhHlBNIc&sig=Vp1PhWGS8iOGj_QxvmWSVH732Cw#v=onepage&q=biology%20of%20the%20potato%20psyllid&f=false

Wen, A., I. Malik, V. Y. Alvarado, J. S. Pasche, X. Wang, H. B. Scholthof, T. E. Mirkov, C. M. Rush, and N. C. Gudmestad. 2009. Detection, distribution, and genetic variability of "*Candidatus* Liberibacter" species Associated with Zebra Complex disease of potato in North America. Plant Disease, 93(11):1102-1115.

Chapter 2

The Mitochondrial Genome of the Potato Psyllid *(Bactericera cockerelli* Sulc.) and Differences among Potato Psyllid Populations of the United States

Abstract

Zebra Chip of potato is caused by the phyto-pathogen *Candidatus* Liberibacter solanacearum (Lso) and is transmitted by the potato psyllid (*Bactericera cockerelli* Sulc.). While the potato psyllid is native to the United States and northern Mexico, Lso range has expanded north and south from over the past ten years. Historically, population-level studies have used the mitochondrial gene Cytochrome Oxidase I (COI) and applied molecular biology techniques to understand the population genetics of the species. Unfortunately, most published phylogenies are based on a few single nucleotide polymorphisms (SNPs) within this single gene sequence. In this study, the nearly complete mitochondrial genome for several potato psyllid populations from North America: Texas, Nebraska, California, Washington, northwest (Washington) and southwest (Arizona), were sequenced using next generation sequencing technology. The gene content included: 2 ribosomal RNAs (16S and 12S rRNAs); the 22 tRNAs and the genes that encode for the 13 proteins: COI-III, Cytb, ND1-6, ND4L, ATP6, and ATP8. Through analysis of nearly complete mitochondrial genome, potato psyllid populations from Texas, Nebraska (central biotype) and southwestern haplotype were determined to be most closely related. Divergent from this clade were the Washington and California populations (western biotype). The northwestern population seems to be a new population that has arisen from the western biotype. The use of the whole mitochondrial genome, instead of one gene, improves the understanding of the

relationships among the different potato psyllid populations as well as with other closely related species by providing greater depth.

Introduction

The potato psyllid (*Bactericera cockerelli* sulc.) is a phloem feeder (Order: Hemiptera) of economic importance because it affects the production of many plants including potatoes. This insect is considered an important agricultural pest because the ability that has to transmit a phyto-pathogen, *Candidatus* Liberibacter solanacearum, the causal agent of zebra chip disease in potato (Munyaneza 2010). Although, the potato psyllid is native to the United States and northern Mexico (Buchman et al. 2012; Nachappa et al. 2012), its geographical range has increased, including the northwestern and western region of the US, some countries in Central America (Guatemala, El Salvador, Honduras and Nicaragua) and New Zealand (Munyaneza 2012; Nachappa et al. 2012; Munyaneza 2007; Secor and Rivera-Varas 2004).

Genetic diversity studies have provided useful insights for understanding patterns of variation of individuals and populations by using genetic markers such as mitochondrial genes (Ekblom and Wolf 2014; Allendorf et al. 2013). Numerous studies, across many taxa, have used the mitochondrial gene Cytochrome Oxidase I (COI) to identify genetic relationships between species, characterize variation, and classify individuals into strains, subspecific clades or haplotypes (Liu et al. 2006). In the case of the potato psyllid, COI sequence analysis has identified two potato psyllid biotypes, which distinguishes insect populations that demonstrate biological and phenological differences from morphologically identical forms (Gordh and Headrick 2011). These biotypes were identified by a single nucleotide polymorphism (SNP) within an amplified COI fragment of 544 bp long (Liu et al. 2006). Chapman et al. (2012), separated the two potato psyllid biotypes: central and western, by using melt temperature analysis and Sybr Green qrt-PCR, reducing the cost of analysis and providing a new platform for phylogenetic studies. Also, high resolution melting analysis and DNA sequencing data of COI

gene has determined the existence of potato psyllid haplotypes, populations that show differences on a single chromosome or a mitochondrial DNA gene (Allendorf et al. 2013), that correlate to the central, western, northwestern and southwestern geographical regions of the United States (Swisher et al. 2012). However, the use of a single gene has limited the correct genetic differentiation of the species and has proved to be of little help. For example, Powell et al. (2012) reported an amplified 3,025 bp fragment of the mitochondria genome containing part of the Cytochrome B (CytB) gene, the complete NADH Dehydrogenase subunit 1 (NAD1) and the complete large subunit rRNA sequence. The results of this analysis revealed that from all seven potato psyllid populations evaluated, within this specific amplicon, they were 98% similar; demonstrating how poorly this amplicon is for genetic variation, as well as how conserved a gene can be (Powell et al. 2012).

The mitochondrial (mt) genome is one of the genomic resources mostly used for systematic entomology (Cameron 2014). The animal mitochondrial DNA (mtDNA) is abundant in the tissues (Li et al. 2012), is relatively small, haploid and generally maternally inherited (Allendorf et al. 2013). MtDNA also, presents a fast rate of evolution (Li et al. 2012), and it usually does not undergoe recombination; thus, it is useful for reconstruction of phylogenies (Allendorf et al. 2013). Unlike nuclear DNA, the historical genealogical record is not shuffled by recombination (Allendorf et al. 2013) and the population size is smaller than nuclear genomes, which results in a shorter expected coalescence time for mtDNA and a higher probability that the mt tree will accurately reflect the species phylogeny (Springer et al. 2001). In insects, mtDNA is a small double-stranded circular molecule of usually 14-20 kb in length (Li et al. 2011) that encodes 37 genes including 13 proteins: COI-III, Cytb, ND1-6, ND4L, ATP6, and ATP8, two ribosomal RNAs (16S and 12S rRNA), and 22 tRNAs (Cameron 2014; Friedrich and Muqim 2003; Chai et al. 2012; Cameron and Whiting 2008). Additionally, the insect mitogenome has at least one sequence known as the A+T-rich region, a major non-coding region, which plays an important role and contains the initiation sites for transcription and replication of the genome

(Chai et al. 2012). The study of more than one mitochondrial gene can give better information for systematic approaches, as study of family, populations and biogeography can be done (Cameron 2014). In this study, the nearly complete mitochondrial genome for the different potato psyllid populations from North America: central, western, northwestern and southwestern, were assembled, annotated, and analyzed by DNA sequencing data to compare and determine genotypic variations among these populations.

Materials and Methods

DNA extraction and sequencing

Potato psyllid adults from reared colonies belonging to the different geographical regions of the United States: central biotype (Texas and Nebraska), western biotype (California and Washington), northwestern haplotype (Washington) and southwestern haplotype (Arizona) were shipped to the laboratory and stored at 95% ethanol, if still alive, and at -20°C prior to be processed. Five adult potato psyllids from each different region were processed, first for mitochondrial isolation and then, for DNA extraction. For an efficient separation and high yield of mitochondria, the Qproteome Mitochondria Isolation Kit (Qiagen) was used, following the supplementary protocol. After isolation of the whole mitochondria, mtDNA was extracted from the mitochondrial fraction, which was treated as a bacteria culture using the DNeasy blood and tissue kit (Qiagen). The mitochondria DNA extractions of all six potato psyllids populations were stored at -20°C, before sequencing, which was done using next generation sequence: MiSeq (Illumina's sequencing – Research and Testing Laboratory, Lubock, TX).

Illumina MiSeq is suitable for this type of research because it reduces costs and run times. For sequencing, the technique uses fluorescently labeled nucleotides and arranges the templates into clusters to facilitate and ensure the high stability of surface-bound template and low non-specific binding of the fluorescently labeled nucleotides. During each sequencing cycle, a single labeled deoxynucleoside triphosphate (dNTP) is added to the nucleic acid chain where

the nucleotide label serves as a terminator for polymerization, which results in an accurately base by base sequencing that eliminates sequence-context specific errors (Illumina Inc. 2010). For the potato psyllid mitochondrial analysis, an average of 1,804,810 sequences reads resulted from the sequencing.

Sequence assembly

The assembly of the mitochondria sequences of the different potato psyllid populations was done using Geneious version 7.1.7 (Biomatters Limited 2014). All raw sequence files were reference assembled (aligned to a known template) and de novo assembled (assembly of short reads to create a full length sequence) into contigs (Figure 2.1). The reference assembly (Figure 2.1A) was performed using the mitochondrial genome of the wolfberry psyllid, *Paratrioza sinica* (NCBI reference sequence: NC_024577.1), with a length of 14,863 bp. The resulting "used reads" for each sample were de novo assembled and then the de novo assembled contigs were reference assembled to *P. sinica* mitochondrial genome. Also, a de novo assembly (Figure 2.1B) of the raw reads for each sample was done, in which a Texas sample contig became in the "new reference" for assembling the remaining samples. All raw reads from the Nebraska, Washington, California, northwestern and southwestern samples were reference assembled to the Texas consensus sequence. A de novo assembly was performed for each of the resulting used reads of each sample, and the de novo assembled contigs of each sample, together with the Texas consensus sequence, were then reference assembled to *P. sinica* to assist in the correct assembly of the new contigs. The contigs from the Nebraska and Washington populations were concatenated, and the nucleotide "N" was added in those places where gaps were found in the sequences. All processes done using Genious version 7.1.7 were run following the default settings of the program. All sequences assembled were then annotated.

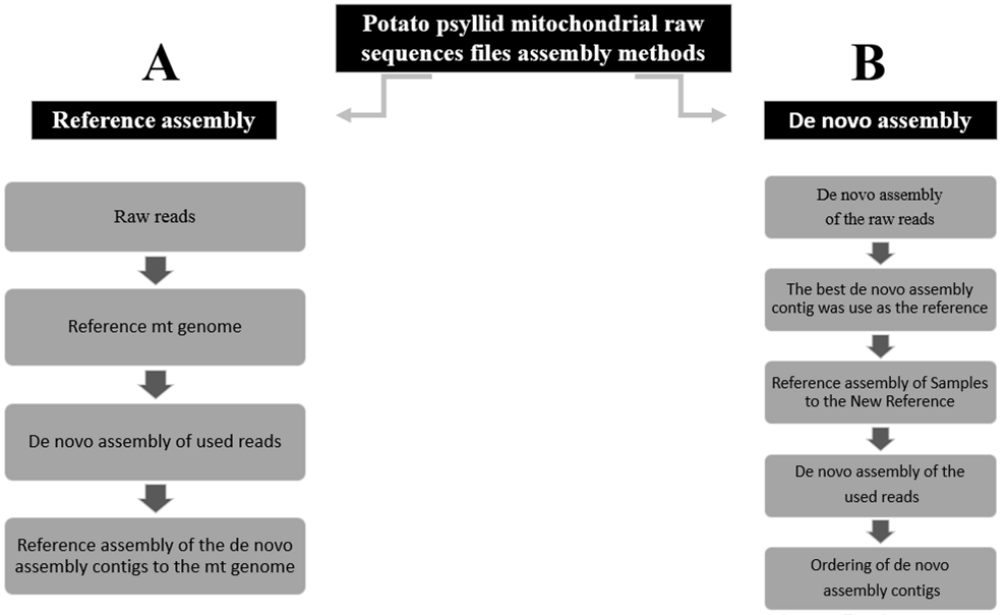


Figure 2.1. Flow chart of the methods and steps followed for assembling the sequencing data: A. Reference assembly, and B. De novo assembly of raw reads.

Annotation of the mitochondrial genome

Protein-coding genes (PCG), ribosomal RNA (rRNA) genes and the transfer RNA (tRNA) genes were identified using the web server MITOS which automatic annotates metazoan mitochondrial genomes (Bernt et al. 2013). MITOS is the most advanced free pipeline available that annotates PCGs by using BLAST searches and covariance methods for annotation of the RNAs (Cameron 2014; Bernt et al. 2013). Since MITOS main annotation results showed a missing gene, the tRNAE (gene that codes for glutamic acid; Figure 2.2), other available web services were used to confirm this absence. ARWEN, a tRNA software detection for metazoan mitochondrial sequences (Laslett and Canbäck 2008), which applies covariation models to detect adjacent hairpin loops and predicts the cloverleaf secondary structure of tRNAs (Cameron 2014; Bernt et al. 2013); and DOGMA, an automatic annotation for annotating plant chloroplast and animal mitochondrial genomes (Wyman et al. 2004), which uses tRNAscan-SE for the identification of tRNAs genes (Bernt et al. 2013).

Furthermore, a more precise identification of the protein-coding genes (PCG) was done by finding open reading frames (ORF) between the tRNAs, using the ORF Finder analysis tool from NCBI and applying the "Invertebrate Mitochondrial" genetic code. Each ORF that was found was then searched against the sequence database using BLAST (NCBI) to verify that the potential code corresponded to the PCG of interest. When the protein-coding regions were predicted, then the start and stop codon for each PCG was identified. This process was done for each of the six potato psyllid populations. Once the tRNAs and the PCGs were predicted, the rRNAs annotation was confirmed by using BLAST, and the putative control region (A+T-rich region) was annotated in between the region of rrnS and the trnI.

Phylogenetic analysis of the mitochondrial genomes

The assembled sequences of all six potato psyllid populations were used to carry out a phylogenetic analysis, and the wolfberry psyllid, *P. sinica* (Hemiptera: Psyllidae), mitochondrial genome was selected as the outgroup. Prior the phylogenetic inference, all potato psyllid and *P. sinica* mt genomes were aligned with MUSCLE v3.8.31 (Edgar 2004) and the resulting file was converted to Phylip4 using Readseq. Model selection was done with jModelTest2 (Darriba et al. 2012; Guindon and Gascuel 2003) with default settings, for Maximum Likelihood analysis and Bayesian inference, respectively. According to Akaike information criterion (AIC) and Bayesian information criterion (BIC), the best-fit model of nucleotide substitution for analysis of the data was the TrN+I model. PhyML 20120412 was employed for the phylogenetic inference, following the default parameters under the TrN+I model. The node support values were assessed by bootstrap resampling calculated using 100 replicates. The resulting tree was visualized in FigTree v1.4.2 (Rambaut 2007).

Results

Genome assembly, composition and organization

The mt genome was assembled by following the de novo assembly approach (Figure 2.1), given that the resulting contigs from the reference assembly only covered two small sections of the entire mitochondrial genome. With the de novo assembly, only one sample, a resulting contig from the Texas sample assembled as a sequence long enough to be called a genome, with a consensus coverage of 664X. Therefore, this contig was used as the new reference for assembling the remaining populations.

The mitochondrial genomes of the six potato psyllid populations exhibit the expected circular DNA molecule (Figure 2.3). The size of the mitochondrial genomes for all six populations was on average 14,523 bp, resulting the Texas population in a genome of 14,536 bp long; the Nebraska population is 14,390 bp long; the California and Northwestern population are 14,492 bp long; Washington resulted in a genome of 14,457 bp, and the Southwestern population had a total of 14,776 bp, being the longest genome of all the potato psyllid populations. This variation in length is mostly due to assembly errors. The gene content observed in the potato psyllid mitochondrial genome of all six potato psyllid populations were arranged as in the ancestral insect mitochondrial genome (Figure 2.2 and 2.3), including: the two ribosomal RNAs (16S and 12S rRNA); the genes that encodes for $tRNA^{alanine}$, $tRNA^{cysteine}$, $tRNA^{aspartic\ acid}$, $tRNA^{phenylalanine}$, $tRNA^{glycine}$, $tRNA^{histidine}$, $tRNA^{isoleucine}$, $tRNA^{lysine}$, $tRNA^{leucine1}$, $tRNA^{leucine2}$, $tRNA^{methionine}$, $tRNA^{aspargine}$, $tRNA^{proline}$, $tRNA^{glutamine}$, $tRNA^{arginine}$, $tRNA^{serine1}$, $tRNA^{serine2}$, $tRNA^{threonine}$, $tRNA^{valine}$, $tRNA^{tryptophan}$, $tRNA^{tyrosine}$ and $tRNA^{glutamic\ acid}$. As well as, the 13 protein coding genes: atp6, atp8, cob, cox1, cox2, cox3, nad1 nad2, nad3, nad4, nad4L, nad5 and nad6; and the non-coding region (putative control region). The orientation of the genes within the genome is the same observed in the ancestral insect mitochondrial genome (Figure 2.2).

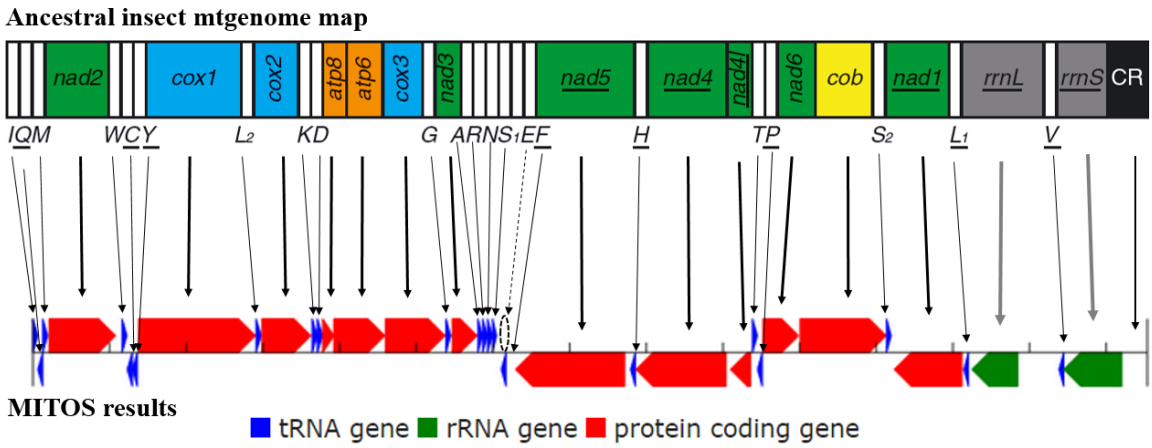
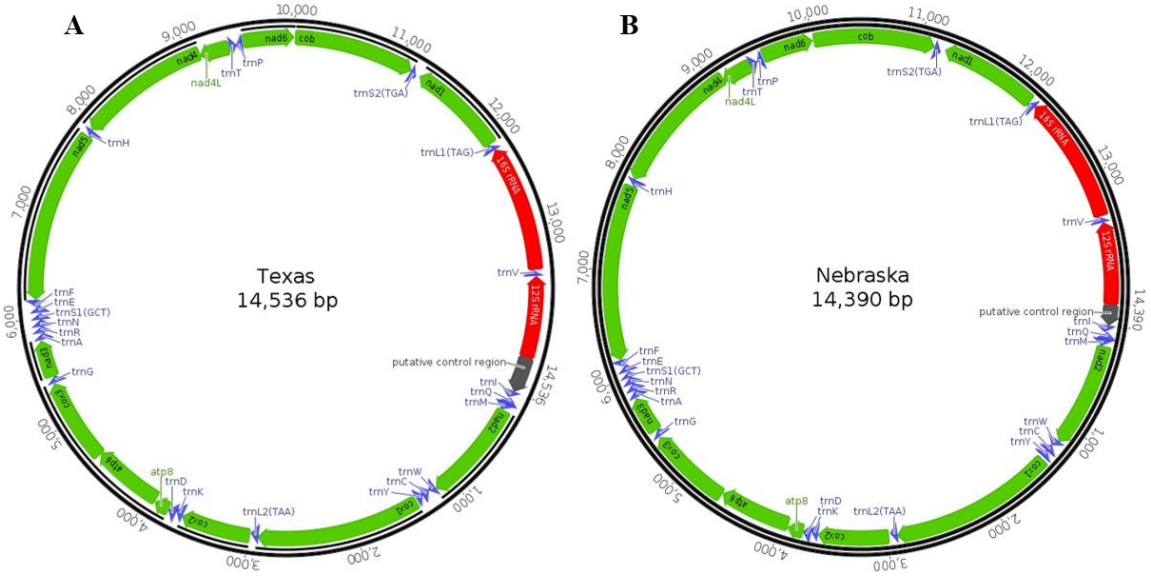


Figure 2.2. Annotation results and arrangements of the genes in the potato psyllid (B. cockerelli Sulc) populations. MITOS annotation results and gene arrangement of the six potato psyllid populations: Texas and Nebraska (central biotype), Washington and California (western biotype), Northwestern and Southwestern haplotypes, as it is observed in the ancestral insect mitochondrial genome (picture of the ancestral insect mtgenome map was adapted from Cameron 2014). In the ancestral insect mitochondrial genome map the genes that are underline represent a reverse orientation of the gene. The dash arrow and circle show were the trnaE gene should appear in MITOS results.



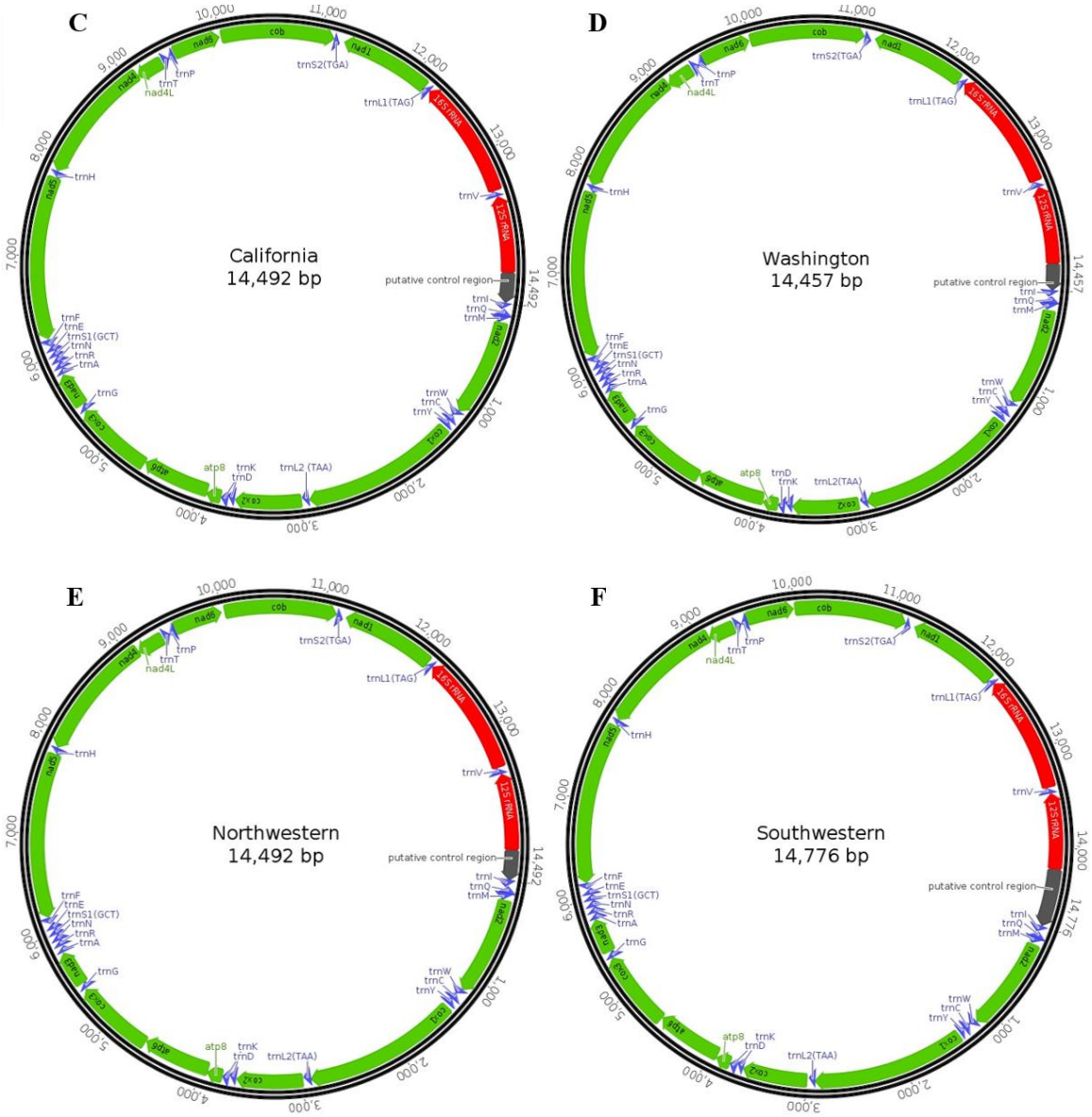


Figure 2.3. Circular DNA molecules of the six potato psyllid (B. cockerelli Sulc) populations. A. Texas, and B. Nebraska (central biotype), C. California, and D. Washington (western biotype), E. Northwestern haplotype, and F. Southwestern haplotype. The PCGs are shown in green, the tRNAs genes in blue, the rRNAs in red, and the control region in gray. The arrows indicate the direction of transcription: forward or reverse.

Overall, for most of the genes, all potato psyllid populations share the same start/stop position within the genome (Tables 2.1.1 – 2.1.6), as well as the length of the gene. In some cases, this variation in gene length is by one or two base pairs. The northwestern haplotype (Table 2.1.5) was the exception, with a cytochrome B gene (cob) which was 21 bp shorter than the other populations. Also, this haplotype exhibit more genes that differ in the start and stop positions of the genes, but the lengths of the genes remain closer or the same to the genes of the other populations. Regarding the PCGs annotations, all six populations show an overlap between the Cox1 and the trnL2 ($tRNA^{leucine2}$) gene, between atp8 and atp6, and between nad4 and nad4L. The only overlap seen in the tRNAs genes is with trnE ($tRNA^{glutamic\ acid}$) and trnF ($tRNA^{phenylalanine}$).

The composition of nucleotides for all six potato psyllid populations was on average 26.3% GC and 73.6% AT. The PCGs start codons (Tables 2.1.1 – 2.1.6) were predominantly an ATN codon: five genes: nad2, cox2, nad6, cob and nad1, start with ATA, four genes: cox1, atp6, cox3 and nad4 start with ATG, nad3 with ATT, and atp8 with ATC. The remaining, nad5 and nad4L started with TTG. The complete termination codons (Tables 2.1.1 – 2.1.6) included: TAA, which was present in seven genes, and TAG, present in only one gene. The remaining PCGs have a T as an incomplete termination codon.

Table 2.1.1. Mitochondrial genome organization of the Texas (central biotype) potato psyllid (*B. cockerelli* Sulc) population

Gene	Strand	Location (bp)	Size (bp)	Start Codon	Stop Codon	Anticodon
trnI	F	1 – 66	66			GAU
trnQ	R	64 – 129	66			UUG
trnM	F	129 -193	65			CAU
nad2	F	197 – 1163	967	ATA	T	
trnW	F	1164 - 1227	64			UCA
trnC	R	1231 - 1297	67			GCA
trnY	R	1298 - 1361	64			GUA
cox1	F	1380 - 2915	1536	ATG	TAA	
trnL2 (TAA)	F	2911 - 2978	68			UAA
cox2	F	2988 - 3641	654	ATA	T	
trnK	F	3643 - 3712	70			CUU
trnD	F	3711 - 3779	69			GUC
atp8	F	3780 - 3929	150	ATC	TAA	
atp6	F	3926 - 4600	675	ATG	TAA	
cox3	F	4600 - 5382	783	ATG	TAA	
trnG	F	5389 - 5449	61			UCC
nad3	F	5450 - 5800	351	ATT	TAA	
trnA	F	5805 - 5865	61			UGC
trnR	F	5867 - 5929	63			UCG
trnN	F	5930 - 5995	66			GUU
trnS1(GCT)	F	5996 – 6049	54			*GC*U
trnE	F	6056 – 6121	66			UUC
trnF	R	6110 – 6172	63			GAA
nad5	R	6173 – 7796	1624	TTG	T	
trnH	R	7797 – 7860	64			GUG
nad4	R	7861 – 9104	1244	ATG	T	
nad4l	R	9097 – 9384	288	TTG	TAG	
trnT	F	9386 – 9451	66			UGU
trnP	R	9452 – 9514	63			UGG
nad6	F	9517 – 9999	483	ATA	TAA	
Cob	F	10008 - 11133	1126	ATA	T	
trnS2(TGA)	F	11134 - 11196	63			UGA
nad1	R	11228 - 12142	915	ATA	TAA	
trnL1(TAG)	R	12143 - 12204	62			UAG
rrnL	R	12205 - 13384	1180			
trnV	R	13385 - 13447	63			UAC
rrnS	R	13448 - 14203	756			
control region		14204 - 14536	333			

Table 2.1.2. Mitochondrial genome organization of the Nebraska (central biotype) potato psyllid (*B. cockerelli* Sulc) population

Gene	Strand	Location (bp)	Size (bp)	Start Codon	Stop Codon	Anticodon
trnI	F	1 – 66	66			GAU
trnQ	R	64 – 129	66			UUG
trnM	F	129 -193	65			CAU
nad2	F	197 – 1163	967	ATA	T	
trnW	F	1164 – 1227	64			UCA
trnC	R	1231 – 1297	67			GCA
trnY	R	1298 – 1361	64			GUA
cox1	F	1380 – 2915	1536	ATG	TAA	
trnL2 (TAA)	F	2911 - 2978	68			UAA
cox2	F	2988 – 3641	654	ATA	T	
trnK	F	3643 – 3712	70			CUU
trnD	F	3711 – 3779	69			GUC
atp8	F	3780 – 3929	150	ATC	TAA	
atp6	F	3926 – 4600	675	ATG	TAA	
cox3	F	4600 – 5382	783	ATG	TAA	
trnG	F	5389 – 5449	61			UCC
nad3	F	5450 – 5800	351	ATT	TAA	
trnA	F	5805 – 5865	61			UGC
trnR	F	5867 – 5929	63			UCG
trnN	F	5930 – 5995	66			GUU
trnS1(GCT)	F	5996 – 6049	54			GCU
trnE	F	6056 – 6121	66			UUC
trnF	R	6110 – 6172	63			GAA
nad5	R	6173 – 7796	1624	TTG	T	
trnH	R	7797 – 7860	64			GUG
nad4	R	7861 – 9104	1244	ATG	T	
nad4l	R	9097 – 9384	288	TTG	TAG	
trnT	F	9386 – 9451	66			UGU
trnP	R	9452 – 9514	63			UGG
nad6	F	9517 – 9999	483	ATA	TAA	
Cob	F	10008 - 11133	1126	ATA	T	
trnS2(TGA)	F	11134 - 11196	63			UGA
nad1	R	11228 - 12142	915	ATA	TAA	
trnL1(TAG)	R	12143 - 12204	62			UAG
rrnL	R	12205 - 13385	1181			
trnV	R	13386 - 13448	63			UAC
rrnS	R	13449 - 14204	756			
control region		14205 - 14390	186			

Table 2.1.3. Mitochondrial genome organization of the California (western biotype) potato psyllid (*B. cockerelli* Sulc) population

Gene	Strand	Location (bp)	Size (bp)	Start Codon	Stop Codon	Anticodon
trnI	F	1 – 66	66			GAU
trnQ	R	64 – 129	66			UUG
trnM	F	129 -193	65			CAU
nad2	F	197 – 1163	967	ATA	T	
trnW	F	1164 – 1227	64			UCA
trnC	R	1231 – 1297	67			GCA
trnY	R	1298 – 1361	64			GUA
cox1	F	1380 – 2915	1536	ATG	TAA	
trnL2 (TAA)	F	2911 - 2978	68			UAA
cox2	F	2988 – 3641	654	ATA	T	
trnK	F	3643 – 3712	70			CUU
trnD	F	3711 – 3779	69			GUC
atp8	F	3780 – 3929	150	ATC	TAA	
atp6	F	3926 – 4600	675	ATG	TAA	
cox3	F	4600 – 5382	783	ATG	TAA	
trnG	F	5389 – 5449	61			UCC
nad3	F	5450 – 5800	351	ATT	TAA	
trnA	F	5805 – 5865	61			UGC
trnR	F	5867 – 5929	63			UCG
trnN	F	5930 – 5995	66			GUU
trnS1(GCT)	F	5996 – 6049	54			GCU
trnE	F	6053 – 6118	66			UUC
trnF	R	6107 – 6169	63			GAA
nad5	R	6173 – 7796	1624	TTG	T	
trnH	R	7794 – 7857	64			GUG
nad4	R	7861 – 9103	1243	ATG	T	
nad4l	R	9097 – 9372	276	TTG	TAG	
trnT	F	9383 – 9448	66			UGU
trnP	R	9449 - 9511	63			UGG
nad6	F	9514 - 9996	483	ATA	TAA	
Cob	F	10005 - 11130	1126	ATA	T	
trnS2(TGA)	F	11131 - 11193	63			UGA
nad1	R	11228 - 12142	915	ATA	TAA	
trnL1(TAG)	R	12140 - 12201	62			UAG
rrnL	R	12202 - 13382	1181			
trnV	R	13383 - 13445	63			UAC
rrnS	R	13446 - 14200	755			
control region		14201 - 14492	292			

Table 2.1.4. Mitochondrial genome organization of the Washington (western biotype) potato psyllid (*B. cockerelli* Sulc) population

Gene	Strand	Location (bp)	Size (bp)	Start Codon	Stop Codon	Anticodon
trnI	F	1 – 66	66			GAU
trnQ	R	64 – 129	66			UUG
trnM	F	129 -193	65			CAU
nad2	F	197 – 1163	967	ATA	T	
trnW	F	1164 – 1227	64			UCA
trnC	R	1231 – 1297	67			GCA
trnY	R	1298 – 1361	64			GUA
cox1	F	1380 – 2915	1536	ATG	TAA	
trnL2 (TAA)	F	2911 - 2978	68			UAA
cox2	F	2988 – 3642	655	ATA	T	
trnK	F	3643 – 3712	70			CUU
trnD	F	3711 – 3779	69			GUC
atp8	F	3780 – 3932	153	ATC	TAA	
atp6	F	3926 – 4600	675	ATG	TAA	
cox3	F	4600 – 5382	783	ATG	TAA	
trnG	F	5389 – 5449	61			UCC
nad3	F	5450 – 5800	351	ATT	TAA	
trnA	F	5805 – 5865	61			UGC
trnR	F	5867 – 5929	63			UCG
trnN	F	5930 – 5995	66			GUU
trnS1(GCT)	F	5996 – 6049	54			GCU
trnE	F	6053 – 6118	66			UUC
trnF	R	6107 – 6169	63			GAA
nad5	R	6170 – 7793	1624	TTG	T	
trnH	R	7794 – 7857	64			GUG
nad4	R	7858 – 9100	1243	ATG	T	
nad4l	R	9094 – 9369	276	TTG	TAG	
trnT	F	9383 – 9448	60			UGU
trnP	R	9449 – 9511	63			UGG
nad6	F	9515 – 9996	483	ATA	TAA	
Cob	F	10005 - 11130	1126	ATA	T	
trnS2(TGA)	F	11131 - 11193	63			UGA
nad1	R	11225 - 12139	915	ATA	TAA	
trnL1(TAG)	R	12140 - 12201	62			UAG
rrnL	R	12202 - 13383	1182			
trnV	R	13384 - 13446	63			UAC
rrnS	R	13447 - 14201	755			
control region		14202 - 14457	256			

Table 2.1.5. Mitochondrial genome organization of the Northwestern potato psyllid (*B. cockerelli* Sulc) haplotype

Gene	Strand	Location (bp)	Size (bp)	Start Codon	Stop Codon	Anticodon
trnI	F	1 – 66	66			GAU
trnQ	R	64 – 129	66			UUG
trnM	F	129 -193	65			CAU
nad2	F	197 - 1163	967	ATA	T	
trnW	F	1164 - 1227	64			UCA
trnC	R	1231 - 1297	67			GCA
trnY	R	1298 - 1361	64			GUA
cox1	F	1385 - 2920	1536	ATG	TAA	
trnL2 (TAA)	F	2916 - 2983	68			UAA
cox2	F	2993 - 3647	655	ATA	T	
trnK	F	3648 - 3717	70			CUU
trnD	F	3716 - 3783	68			GUC
atp8	F	3784 - 3933	150	ATC	TAA	
atp6	F	3930 - 4604	675	ATG	TAA	
cox3	F	4604 - 5386	783	ATG	TAA	
TrnG	F	5391 - 5451	61			UCC
nad3	F	5452 - 5802	351	ATT	TAA	
trnA	F	5807 - 5867	61			UGC
trnR	F	5869 - 5930	62			UCG
trnN	F	5931 - 5996	66			GUU
trnS1(GCT)	F	5997 - 6050	54			GCU
TrnE	F	6054 - 6119	66			UUC
TrnF	R	6108 - 6170	63			GAA
nad5	R	6171 - 7794	1624	TTG	T	
trnH	R	7795 - 7858	64			GUG
nad4	R	7859 - 9101	1243	ATG	T	
nad4l	R	9095 - 9370	276	TTG	TAG	
trnT	F	9384 - 9449	66			UGU
trnP	R	9450 - 9513	64			UGG
nad6	F	9516 - 9998	483	ATA	TAA	
Cob	F	10028 - 11132	1105	ATA	T	
trnS2(TGA)	F	11133 - 11195	63			UGA
nad1	R	11227 – 12141	915	ATA	TAA	
trnL1(TAG)	R	12142 – 12203	62			UAG
rrnL	R	12204 – 13385	1182			
trnV	R	13386 – 13448	63			UAC
rrnS	R	13449 – 14203	755			
control region		14204 – 14492	289			

Table 2.1.6. Mitochondrial genome organization of the Southwestern potato psyllid (*B. cockerelli* Sulc) haplotype

Gene	Strand	Location (bp)	Size (bp)	Start Codon	Stop Codon	Anticodon
trnI	F	1 – 66	66			GAU
trnQ	R	64 – 129	66			UUG
trnM	F	129 -193	65			CAU
nad2	F	197 – 1163	967	ATA	T	
trnW	F	1164 – 1227	64			UCA
trnC	R	1231 – 1297	67			GCA
trnY	R	1298 – 1361	64			GUA
cox1	F	1380 – 2915	1536	ATG	TAA	
trnL2 (TAA)	F	2911 - 2978	68			UAA
cox2	F	2988 – 3642	655	ATA	T	
trnK	F	3643 – 3712	70			CUU
trnD	F	3711 – 3779	69			GUC
atp8	F	3780 – 3929	150	ATC	TAA	
atp6	F	3926 – 4600	675	ATG	TAA	
cox3	F	4600 – 5382	783	ATG	TAA	
trnG	F	5389 – 5449	61			UCC
nad3	F	5450 – 5800	351	ATT	TAA	
trnA	F	5805 – 5865	61			UGC
trnR	F	5867 – 5929	63			UCG
trnN	F	5930 – 5995	66			GUU
trnS1(GCT)	F	5996 – 6049	54			GCU
trnE	F	6055 – 6120	66			UUC
trnF	R	6109 – 6172	63			GAA
nad5	R	6172 – 7795	1624	TTG	T	
trnH	R	7796 – 7859	64			GUG
nad4	R	7860 – 9102	1243	ATG	T	
nad4l	R	9096 – 9371	276	TTG	TAG	
trnT	F	9385 – 9450	66			UGU
trnP	R	9451 – 9513	63			UGG
nad6	F	9516 – 9998	483	ATA	TAA	
Cob	F	10007 - 11132	1126	ATA	T	
trnS2(TGA)	F	11133 - 11195	63			UGA
nad1	R	11227 - 12141	915	ATA	TAA	
trnL1(TAG)	R	12142 - 12203	62			UAG
rrnL	R	12204 - 13383	1180			
trnV	R	13384 - 13446	63			UAC
rrnS	R	13447 - 14202	756			
control region		14203 - 14776	574			

The 22 transfer RNA genes (Figure 2.4) for all six potato psyllid populations fold into the typical cloverleaf structure, except for trnS1 ($tRNA^{Serine1}$) that lacks the dihydrouridine (DHU) arm. All tRNA genes ranged from 54 to 70 bp, and the anticodons for each of the 22 tRNA genes were alike for all the different potato psyllid populations (Tables 2.1.1-2.1.6). The northwestern haplotype show differences (Figure 2.4) in some of the nucleotides that form the secondary structure of trnR ($tRNA^{arginine}$), trnT ($tRNA^{threonine}$), trnL1 ($Trna^{leucine1}$) and trnV ($tRNA^{valine}$) genes.

The ribosomal RNAs, the large (rrnL) and small (rrnS) subunits, in all potato psyllid populations are located after trnL ($tRNA^{leucine1}$), trnV ($tRNA^{valine}$) genes and before the control region, respectively. Depending on the potato psyllid population, the lengths for the rrnL vary from 1180 to 1182 bp, and for the rrnS was either 755 or 756 bp (Tables 2.1.1-2.1.6). The putative control region, for all populations, is located in between the rrnS and trnI- trnQ- trnM gene cluster and a part of this region needs to be sequenced in order to be completed (Tables 2.1.1-2.1.6).

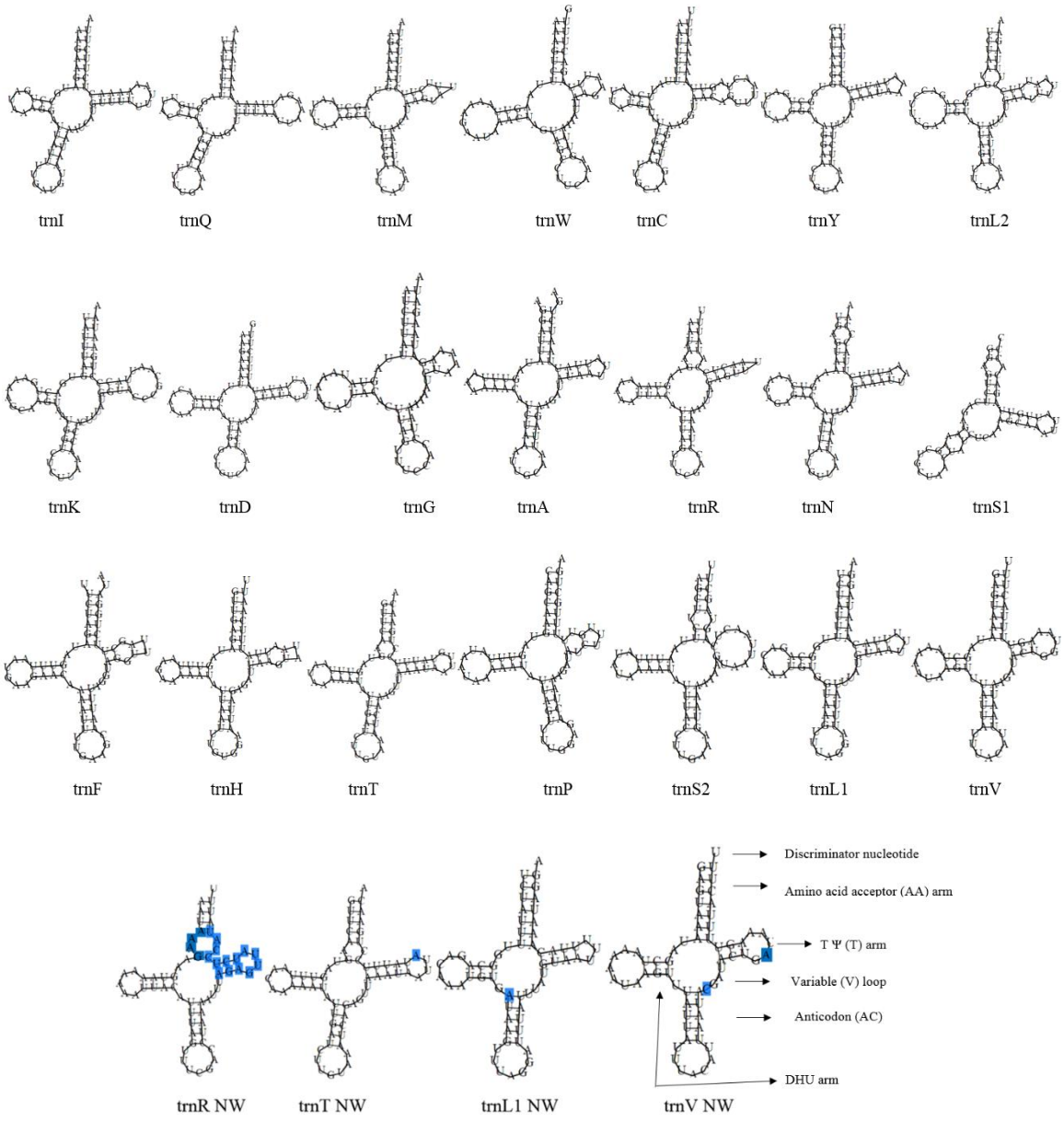


Figure. 2.4. Transfer RNA (tRNA) genes structure. Secondary structures of the tRNA genes, showing the typical cloverleaf structure, of the mitochondrial genomes of Texas, Nebraska, California, Washington, Northwestern and Southwestern potato psyllid populations. *The Northwestern (NW) population show differences in the nucleotides (in blue) that make up the cloverleaf structure in four of the tRNAs: trnR, trnT, trnL1 and trnV.*

Phylogenetic analysis

A phylogenetic tree (Figure 2.5) was built using the mitochondrial genomes of the different potato psyllid populations: central biotype (Texas and Nebraska), western biotype (California and Washington), northwestern haplotype and southwestern haplotype, and *P. sinica* (the wolfberry psyllid) was used as outgroup, because of the close relationship of this organism with the potato psyllid. Both *P. sinica* and *B. cockerelli* are important pest of Solanacea (but different genus), and belong to the same suborder of Hemiptera (*Sternorryncha*), but to different families. Therefore, *P. sinica* was hypothesized to be closely related enough to the potato psyllid, but not as closely related as the other potato psyllid populations are to each other. Also, the branches of the resulting tree were transformed for a better visualization.

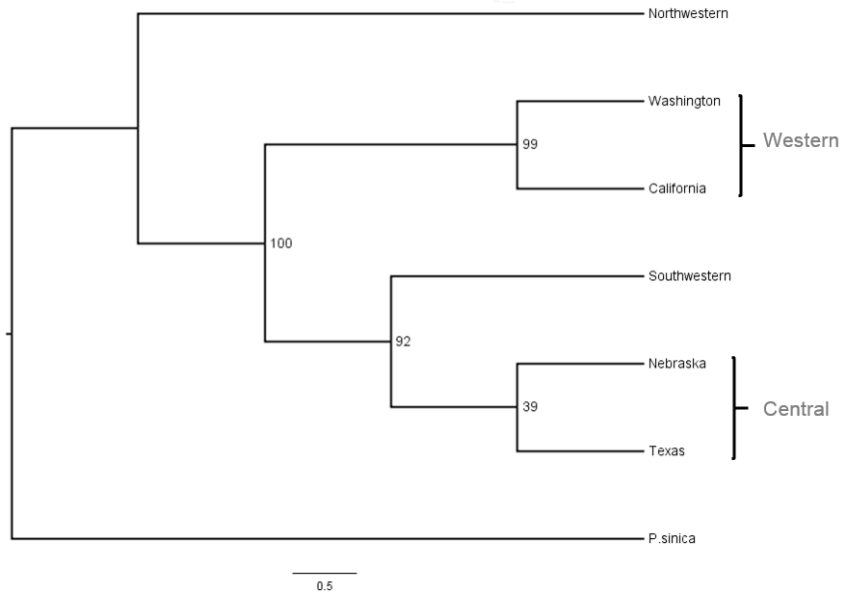


Figure 2.5 Phylogenetic inference of the potato psyllid (B. cockerelli Sulc) populations. A view of the relationships of the different potato psyllids populations: central biotype (Texas and Nebraska), western biotype (California and Washington), northwestern and southwestern haplotype; and Paratrioza sinica, the wolfberry psyllid, used as outgroup. Maximum likelihood and Bayesian inference was analyzed with PhyML 20120412 under the TrN+I model. The node support values were assessed by bootstrap resampling calculated using 100 replicates.

Discussion

Potato psyllid mitochondrial genome organization

The organization of the potato psyllid mtgenomes, its content, and gene order is the same as the one observed in the other two family species mitochondrial genomes: *Paratrioza sinica* and *Pachypsella venusta*, sharing the same gene content and order as the one seen in the ancestral insect mitochondrial genome. Although insect mitochondrial genomes are typically 15-18 kb in size (Cameron 2014), the six potato psyllid mitochondrial genomes were as the other Psyllidae species in the 14 kb size range, *P. sinica* is 14,863 bp and *P. venusta* is 14,711 bp (Zhang et al. 2014). The southwestern haplotype mitochondrial genome was the longest, with a size of 14,776 bp long and the Nebraska mitogenome was the smallest from all potato psyllid genomes, with 14,390 bp long. The observed differences in sizes are mainly due to artifacts of the assembly process. The sizes of the control regions were not as long as in the other Psyllidae species, e.g. *P. sinica*: 700 bp (Zhang et al. 2014) or other Hemiptera species like the assassin bug, *Agriosphodrus dohrni*, which is 1,643 bp long (Li et al. 2011). However, this control region among arthropods is highly variable in size, can be as short as 73 bp or as long as 4.6kb as in *Drosophila melanogaster* (Simon et al. 2006), and between 9-13 kb in bark weevils (Simon et al. 1994). This size difference in the potato psyllid genomes can be improved by resequencing the control region. Although, evidence indicates that in arthropods this region can represent a challenge when trying to amplify, sequence and analyze it, because of its high content of tandem repeats (Simon et al. 2006). This repeated sequence blocks of one or more nucleotides, mainly AT, makes it difficult to design primers. In addition, primers with only AT require low anneal temperatures and tend to be unselective when binding (Simon et al. 2006).

The GC and AT content observed in the mt genomes of the biotypes and haplotypes of the potato psyllid, 26.3% and 73.6%, respectively, is consistent with what it is seen among insect mt genomes. In general, these genomes show a strong base compositional bias were the AT and GC content is unequal, and tend to be rich in AT, not only in the whole mt genome, but also in

individual mt genes (Simon et al. 1994). Also, this high AT content varies across (in termites AT content is 64% and in bees 86.7%), and within orders (AT content in beetles goes from 65% to 78%) (Cameron 2014). Data suggests that insect mt genomes have adapted to AT richness conditions and selection acts to maintain it (Simon et al 1994).

In the case of the tRNA secondary structure, the lack of the dihydrouridine (DHU) arm in the trnS1 gene ($tRNA^{serine1}$) and; therefore, the absence of the potential cloverleaf structure in this gene is a distinctive feature of the metazoan mitochondrial genomes (Cameron 2014; Zhang et al. 2014; Li et al. 2011). Moreover, the expected start and stop codon usage was observed for the 13 PCGs in the potato psyllid biotypes and haplotypes. These start codons were either a Met or Ile start codon (ATG, ATA, ATC or ATT), or either a Lys or Val start codon (TTG). As for the stop codons, complete termination codons (TAA) and partial stop codons (T) were determined. Both star/stop codons seen in the potato psyllid genomes are the ones used among the insect taxa (Cameron 2014).

Regarding the ribosomal RNAs, since the rRNA genes are the most difficult to annotate, the large rRNA (rrnL) has been annotated consistently to occupy every base between the two tRNAs (trnV and trnL1) where is located within the genome (Cameron 2014). For the potato psyllids genomes, the region corresponding to the rrnL was verified using BLAST. The small rRNA (rrnS) annotation was left as the one predicted by MITOS, given that it is easy to know where the gene begins but not where it stops, and MITOS includes a software (Infernal) that provides a more consistent annotation of the rRNAs genes (Cameron 2014).

Phylogenetic relationships among the potato psyllid populations via their mitochondrial genomes

The phylogenetic analysis was consistent with the information already known about the potato psyllids populations, showing that the populations from Texas, Nebraska and southwest are geographically closely related, as well as the populations from Washington and California.

The northwestern population seems to be an isolated population from the western populations. However, populations belonging to the central (Texas and Nebraska) and western (California and Washington) region of the United States have been identified as biotypes because of their biological differences, such as survivorship, growth index and developmental time (Swisher et al. 2013). Whereas, the northwestern and southwestern are considered as haplotypes because, so far, only their genetic differences have been identified by using a single mt gene (Swisher et al. 2013). Hence, by using the whole mt genome of the potato psyllid, the central biotype and southwestern haplotype are not that considerably different. When comparing the central biotype and southwestern haplotype genomes, a difference of 1 bp in the start and stop position of 13 genes (cox2, trnE, trnF, nad4, trnT, trnP, nad6, cob, trnS2, nad1, trnL1, rrnL, trnV and rrnS) exists and difference in length in the nad4L gene was documented. The latter might support the haplotype hypothesis when differentiating this population as a different haplotype by using one mitochondrial gene. However, there are 37 genes that encompass the whole mitochondrial genome, and the differences, in the southwestern haplotype, are not even seen in half of the genes of the genome. Therefore, if no establish consent of how statistically different populations need to be in order to be called a new population, and because the southwestern haplotype clusters together with the other populations from the central biotype, it is probably not correct to say that is a new haplotype. On the other hand, if so, then it is clear that the southwestern haplotype resulted from the central biotype.

Regarding to the northwestern haplotype, this is a population that has been isolated from the others for a long time, which is obviously due to the geography of the Pacific Northwest and Columbia Basin. Also, haplotyping results from archived psyllids collected from the northwestern region have shown that this population has potentially been present in this region longer than the western population (Swisher et al. 2013). In addition, by comparing the gene organization of the potato psyllid mitochondrial genomes, the northwestern haplotype is different. The position where 28 of the genes start and stop within the genome varies, from 1-5 bp or even 21 bp of

difference as it is seen in the cytochrome oxidase B gene (cob), showing a larger genetic variation with the other potato psyllid haplotypes and biotypes. Also, this genetic variation is greatly supported by the large number of single nucleotide polymorphisms (SNPs) within the mt genes of the northwestern population, when compared to the other populations. Furthermore, since the northwestern and western populations are present in the same area, the latter results suggest that the northwestern population is native to this area, while the western population has migrated (Swisher et al. 2013). Moreover, migration of the western population to the Pacific Northwest and Columbia Basin can lead to gene flow between these two populations, a source of genetic variation. Unfortunately, since mtDNA acts as a single locus because it does not undergoes recombination, is not useful for describing genetic population structure within species, perhaps because of its faster rate of introgression. Nevertheless, there can be considerable differences in terms of genetic variation. This genetic variation seen in the northwestern haplotype is mostly due to natural selection and the means to adapt to this climatic zone, as well as a reduction in gene flow. Since, migration of this population might be limited due to the geographic barriers, mating does not occur randomly, and the gene frequencies in the population are altered.

Conclusions

A better understanding of the potato psyllid population dynamics and the proper identification of biotypes and haplotypes of *B. cockerelli* is crucial, especially when talking about important pest crops. The use of a single mt gene, in the past, has revealed differences between populations. With the mtgenome the difference of the existence of a central and western biotypes are supported. In addition, the mt genome is also showing the great genotypic variation found in the northwestern population, and consequently a most likely phenotypic variation. Moreover, in our findings, the southwestern population should not be considered any more as a different population because is not showing any genotypic variation. Thus, only central, western and northwestern should be called biotypes. Additionally, whole mtgenome analysis can be applied to

know how species that are vectors of pathogens, such as the potato psyllid, are related to others and determine if the ones that are not considered a harm, can become in one.

References

Allendorf, F. W., G. Luikart, and S. N. Aitken. 2013. Conservation and the Genetics of Populations. Second edition. Wiley-Blackwell, Malaysia. 602 pp.

Bernt, M., A. Braband, M. Middendorf, B. Misof, O. Rota-Stabelli, and P. F. Stadler. 2013. Bioinformatics methods for the comparative analysis of metazoan mitochondrial genome sequences. Molecular Phylogenetics and Evolution, 69: 320-327.

Bernt M., A. F. Donath, F. Jühling, C. Externbrink, G. Florentz, J. Fritzsch, M. Pütz, P.F. Middendorf, and P. F. Stadler. 2013. MITOS: Improved de novo Metazoan Mitochondrial Genome Annotation. Molecular Phylogenetics and Evolution, 69(2):3f13-319.

Biomatters. 2014. Geneious v7.1.7. www.geneious.com

Buchman, J., T. W. Fischer, V. G. Sengoda, and J. E. Munyaneza. 2012. Zebra Chip Progression: From Inoculation of Potato Plants with Liberibacter to Development of Disease Symptoms in Tubers. Am. J. Pot Res 89:159-168.

Cameron, S. L. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu. Rev. Entomol. 59: 95-117.

Cameron, S. L. 2014. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. Systematic Entomolgy 39: 400-411.

Cameron, S. L., and M. F. Whiting. 2008. The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. Gene 408: 112-123.

Chai, H-N., Y-Z. Du, and B-P Zhai. 2012. Characterization of the Complete Mitochondrial Genomes of *Cnaphalocrocis medinalis* and *Chilo suppressalis* (Lepidoptera: Pyralidae). Int. J. Biol. Sci. 8(4): 561-579.

Chapman, R. I., J. F. Macias-Velasco, A. P. Arp, and Blake R. Bextine. 2012. Using quantitative real time PCR melt curve analysis of partial CO1 sequence for rapid biotype differentiation of *Bactericera cockerelli* (Hemiptera: Triozidae). Southwestern entomologist, 37(4): 2012.

Darriba, D., G. L. Taoba, R. Doallo, D. Posada. 2012. jModelTest2: more models, new heuristics and parallel computing. Nature Methods 9(8): 772.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32(5): 1792- 1797.

Ekblom, R., and J. B. W. Wolf. 2014. A field guide to whole-genome sequencing, assembly and annotation. Evolutionary Applications 7(9): 1026-1042.

Friedrich, M., and N. Muqim. 2003. Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castanaeum*. Molecular Phylogenetics and Evolution 26:502-512.

Gordh, G., and D. Headrick. 2011. A dictionary of Entomology. Second edition, CSIRO Publishing, 1526 pp.

Guindon, S., and O. Gascuel. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood". Systematic Biology 52: 696-704.

Illumina Inc. 2010. Illumina Sequencing Technology: highest data accuracy, simple workflow, and a broad range of applications. Technology Spotlight: Ilumina Sequencing. Illumina, Inc. Retrieved from: http://res.illumina.com/documents/products/techspotlights/techspotlight_sequencing.pdf

Laslett, D., and B. Canbäck. 2008. ARWEN, a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24:172-175.

Li, H., H. Liu, A. Shi, P. Stys, X. Zhou, and W. Cai. 2012. The Complete Mitochondrial Genome and Novel Gene Arrangement of the Unique-Headed Bug *Stenopirates* sp. (Hemiptera: Enicocephalidae). PLoS ONE, 7(1): e29419. doi:10.1371/journal.pone.0029419

Li, H., J. Gao, H, Liu, H. Liu, A. Liang, X. Zhou, and W. Cai, 2011. The Architecture and Complete Sequence of Mitochondrial Genome of an Assassin Bug *Agriosphodurs dohrni* (Hemiptera: Reduviidae). Int. J. Biol. Sci, 7(6): 792-804.

Liu, D., J. T. Trumble, and R. Stouthamer. 2006. Genetic differentiation between eastern populations and recent introductions of potato psyllid (*Bactericera* cockerelli) into western North America. Entomologia Experimentalis et Applicata 118: 177-183.

Munyaneza, J. 2010. Psyllids as vectors of Emerging Bacterial Diseases of Annual Crops. Southwestern Entomologists, 35(3): 471-477.

Munyaneza, J.E., J.M. Crosslin and J. E. Upton. 2007. Association of *Bactericera* cockerelli (Homoptera: Psyllidae) with "Zebra Chip", a New Potato Disease in Southwestern United States and Mexico. Journal of Economic Entomology 100 (3): 656-663.

Nachappa, P., J. Levy, and C. Tamborindeguy. 2012. Transcriptome analyses of *Bactericera cockerelli* adults in response to "*Candidatus* Liberibacter solanacearum" infection. Mol Genet Genomics 287: 803-817.

Powell, C. M., J. King, and B. R. Bextine. 2012. Cytochrome B Sequences of Potato Psyllids, *Bactericera cockerelli* (Sulc), from North and Central America. Southwestern Entomologist, 37(4): 521-523.

Rambaut, A. 2007. Molecular evolution, phylogenetics and epidemiology. Retrieved from: http://tree.bio.ed.ac.uk/software/

Secor, G. A. and V. V. Rivera-Varas. 2004. Emerging diseases of cultivated potato and their impact on Latin America. Rev. Latinoamericana de la papa (suplemento) 1:1-8.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87(6): 651-701.

Simon, C., T. R. Buckley, F. Frati, J. B. Stewart, and A. T. Beckenbach. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved

polymerase chain reaction primers for animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 37: 545-579.

Springer, M. S., R. W. DeBry, C. Douady, H. M. Amrine, O. Madsen, W. W. de Jong, and M. J. Stanhope. 2001. Mitochondrial Versus Nuclear Gene Sequences in Deep-Level Mammalian Phylogeny Reconstruction. Mol. Biol. Evol. 18(2): 132-143.

Swisher, K.D, J. E. Munyaneza, and J. M. Crosslin. 2013. Temporal and Spatial Analysis of Potato Psyllid Haplotypes in the United States. Environ. Entomol. 42(2): 381-393.

Swisher, K. D., J. E. Munyaneza, and J. M. Crosslin. 2012. High Resolution Melting Analysis of the Cytochrome Oxidase I Gene Identifies Three Haplotypes of the Potato Psyllid in the United States. Environmental Entomology, 41(4): 1019-1028.

Wyman, S.K., R. K. Jansen, and J. L. Boore. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20(17): 3252-3255.

Zhang, Q-L., Z-L. Guo, and M-L. Yuan. 2014. The complete mitochondrial genome of *Poratrioza sinica* (Insecta: Hemiptera: Psyllidae). Mitochondrial DNA: 1-2.

Chapter 3

Comparison of Potato Psyllid Populations Relationship and Identification of Potato Psyllid Populations by using the Whole Mitochondrial Genome and Single Mitochondrial Genes

Abstract

Over the past 10 years, the potato psyllid (*Bactericera cockerelli* sulc.), a Solanaceous pest and vector of the zebra chip disease phyto-pathogen, (*Candidatus* Liberibacter solanacearum), has increased its geographical range. Initially, variation between species was identified based on morphological characters, but molecular traits were included (DNA barcoding) when morphology was not enough to make a differentiation. The cytochrome c oxidase subunit 1 (COI) gene has been set as the DNA barcode region for metazoans for species identification, delimitation and construction of phylogenetic relationships. COI *B. cockerelli* genetic studies have differentiated the potato psyllid biotypes (central and western) and haplotypes (northwestern and southwestern). In this study, the 37 genes (protein coding genes, tRNA genes and rRNA genes) of the mt genome of the potato psyllid haplotypes from central (Texas and Nebraska), western (California and Washington), northwest, and southwest, were used for building phylogenetic trees with each gene, using Geneious tree builder (Geneious 8.1). The 37 inferred trees were assessed by bootstrap values and a double cluster dendrogram was performed for a global visualization of the data presented by each phylogenetic tree. From all 37 genes, only COI, COIII and ATP6 organize the potato psyllid populations similar to whole mt genome phylogeny: Texas, Nebraska (central biotype) and southwestern as one clade; Washington and California populations (western biotype) as a second clade, and northwestern as

an isolated population. Therefore, single Genes highly susceptible to variation (COI, COII and ATP6) can be applied for phylogenetic studies of the potato psyllid.

Introduction

The potato psyllid (*Bactericera cockerelli* sulc.) is a small sap-sucking insect and a major pest of solanaceous crops, affecting mainly the production of potatoes (Buchman et al. 2012; Butler and Trumble 2012; Nachappa et al. 2012; Munyaneza 2010; Munyaneza et al. 2007). This insect is the vector of the phytopathogen, *Candidatus* Liberibacter solanacearum (Lso), which causes the zebra chip disease (Ramírez-Davila et al. 2012; Buchman et al. 2011; Munyaneza 2010). Although, the potato psyllid is native to the United States and northern Mexico (Buchman et al. 2012; Nachappa et al. 2012), its geographical range has increased, including the northwestern and western region of the US, some countries in Central America and New Zealand (Nachappa et al. 2012). Moreover, the plant pathogen (Lso), in the last years, has also been detected on carrots from Finland, Norway, Sweden and the Mediterranean Region, as well as in celery crops from Spain; evidencing that the bacteria has more than one host and more than one vector (Munyaneza 2012). Depending on the region, the bacteria has different psyllid vectors such as *Trioza apicalis* (Finland), *Bactericera trigonica* (Canary Islands and Spain), and *Bactericera cockerelli* (USA, Central America and New Zealand) (Janse 2012; Munyaneza 2012).

Before the advances in molecular techniques, the studies of variation of species were based only on morphological characters. However, the understanding of variability among organisms and populations has increased and provided many more tools for studying phylogenetic relationships by including the variations at the molecular level, which are based on changes of DNA nucleotides sequences of homologous genes (Ajmal et al. 2014). The correct identification of insects is vital, especially for those that are of agricultural importance and need to diagnose pest damage and to accurately determine which control measures are appropriate

(Meeyen et al. 2014; Golick et al. 2013). Morphological identification is the conventional, gold standard method to identify species based on external characteristics by using taxonomic keys. However, this method requires experienced taxonomist, is highly time-consuming, and incomplete identification is often faced when important morphological features are missing or damaged due to poor specimen handling (Chan et al. 2014). In addition, the morphological identification becomes problematic when morphology is not an enough feature to distinguish species, especially when working at the subspecies, cultivars, ecotypes, morphotypes, clones (Ajmal et al. 2014), biotypes and haplotypes level. Also, with the sample, often only parts of the insect are found.

DNA barcoding is a recent molecular-based and bioinformatics tool for identification, which has been widely used to identify biological specimens and assign them to a given species, as well as to investigate species interactions (Ajmal et al. 2014; Bergmann et al. 2013; Collins and Cruickshank 2013; Casiraghi et al. 2010). The method is based on partial mitochondrial DNA sequences between 400-800 base pairs that can be used as genetic markers (Ajmal et al. 2014; Chan et al. 2014; Bergmann et al. 2013; Casiraghi et al. 2010). Early studies used the nuclear internal transcribed spacer 2, cytochrome b oxidase, 12S rRNA and nicotinamide adenine dinucleotide dehydrogenase (NADH). However, the Consortium for the Barcode of Life (CBoL), agreed to use cytochrome c oxidase subunit 1 (COI) gene as a default DNA barcode region for metazoans (Chan et al. 2014; Bergmann et al. 2013; Casiraghi et al. 2010). COI gene gained popularity because of the existence of universal primers to amplify the gene and the ability to provide a higher genetic variation between species than within species (Chan et al. 2014; Casiraghi et al. 2010).

In insects, the mt genome is a circular molecule of usually 14-20 kb in length (Cameron 2014; Li et al. 2011) that encodes for 37 genes: 2 ribosomal RNAs (16S and 12S rRNA); 22 tRNAs and 13 proteins (COI-III, Cytb, ND1-6, ND4L, ATP6, and ATP8) (Cameron 2014; Friedrich and Muqim 2003; Chai et al. 2012; Cameron and Whiting 2008). Mt genomes and mt

genes are the most used sequence data in animal phylogenetics for reconstructing evolutionary relationships among taxa at particular levels of divergence. The high usage of mtDNA is due to its abundance, fast rate of evolution, ease of sequencing, nonrecombination and conserved gene content (Simon et al. 2006; Simon et al. 1994).

Reliable DNA barcodes using COI, have been successfully obtained, especially for arthropods (Bergmann et al. 2013), and have been commonly used for identifying species belonging to the orders: Diptera, Lepidoptera, Ephemenoptera, Hemiptera, Coleoptera and Hymenoptera (Lee and Akimoto 2015). Besides species identification, COI has been used to identify genetic relationships between species, characterize variation, and classify individuals into strains, subspecific clades or haplotypes (Liu et al. 2006). In the family Triozidae, COI has helped to identified two biotypes of the potato psyllid. These biotypes were identified by a single nucleotide polymorphism (SNP) within an amplified COI fragment of 544 bp long (Liu et al. 2006). In other study, Chapman et al. (2012) facilitated the identification of the two potato psyllid biotypes: central and western, by using COI melt temperature analysis and Sybr Green qrt-PCR. Also, high resolution melting analysis and DNA sequencing data of COI gene has determined the existence of potato psyllid haplotypes that correlate to the central, western, northwestern and southwestern geographical regions of the United States (Swisher et al. 2012). Moreover, other mitochondrial genes have been used to study potato psyllid genetic variations. A mitochondria genome fragment of 3,025 bp containing part of the Cytochrome B (CytB) gene, the complete NADH Dehydrogenase subunit 1 (NAD1) and the complete large subunit rRNA sequence, demonstrated 98% similarity among seven potato psyllid populations. The results of this analysis evidenced how poorly this amplicon is for determining genetic variation because of its conserved gene content (Powell et al. 2012).

The mt genes have been used widely in insect molecular systematic studies because they are easy to sequence and sequencing methods were still needed to be improved for sequencing whole mt genomes (Cameron 2014; Simon et al. 1994). Moreover, both whole mt genomes and

single mt genes have been used to understand molecular systematics, population genetics, diagnostics, and molecular evolutionary studies (Cameron 2014). In this study, the genetic content of the potato psyllid mt genome was evaluated for phylogenetic analysis to determine if the use of single mt genes can resolve the questions within the potato psyllid populations in relation to their genotypic variations, or if the 37 genes that encompass the mt genome are needed to fully answer these questions.

Material and methods

DNA extraction and sequencing

Potato psyllid adults from reared colonies belonging to the different geographical regions of the United States: central biotype (Texas and Nebraska), western biotype (California and Washington), northwestern haplotype (Washington) and southwestern haplotype (Arizona) were shipped to the laboratory and stored at 95% ethanol, if still alive, and at -20°C prior to be processed. Mitochondrial isolation of five adult potato psyllids representing each of the different geographical regions was performed using the Qproteome Mitochondria Isolation Kit (Qiagen). Mitochondrial isolation was followed by DNA extraction, using DNeasy blood and tissue kit (Qiagen). The mitochondria DNA extractions were sent for sequencing, and next generation sequence: MiSeq (Illumina's sequencing) was performed. Among the benefits of this sequencing platform are: reduce costs and run times, and the accurately base by base sequencing that eliminates sequence-context specific errors (Illumina Inc. 2010). For the potato psyllid mitochondrial analysis, an average of nearly 2 million sequences reads resulted from the sequencing.

Sequence assembly and annotation

The raw sequences files of the six potato psyllid populations were assembled using Geneious version 7.1.7 (Biomatters Limited 2014) and the default settings of the program were

followed. All raw reads were assembled into contigs by performing a reference assembly (aligned to a known template) and a de novo assembly (assembly of short reads to create a full length sequence). The reference assembly was performed using the mitochondrial genome of the wolfberry psyllid, *Paratrioza sinica* (NCBI reference sequence: NC_024577.1, with accession name KJ650081). The de novo assembly of the raw reads for each population resulted in the use of a Texas sample contig as the "new reference" to reference assembly the remaining populations. The used reads of the reference assembly were then de novo assembled into contigs. The de novo assembled contigs of each sample, together with the Texas consensus sequence, were reference assembled to *P. sinica* to assist in the correct assembly of the new contigs. The contigs from the Nebraska and Washington populations had to be concatenated, and the nucleotide "N" was added in those places where gaps were found in the sequences. All sequences assembled were then annotated using the web server MITOS, as the main tool for determining the mt genes; DOGMA and ARWEN to assist in confirming the presence of the tRNA genes, and ORF Finder analysis tool from NCBI to correctly identify the protein-coding genes (PCG).

Phylogenetic analysis of the mitochondrial genes

The nearly complete mt genomes of all six potato psyllid populations were used to carry out a phylogenetic analysis. The mt genome of the wolf-berry psyllid, *P. sinica*, was selected as outgroup. Before building the trees, a whole genome multiple alignment was constructed by using MAUVE in Geneious (version 8.1) to identify any rearrangements and inversions within the genomes (Biomatters 2015). With the resulting alignment, the aligned blocks corresponding to each mt gene (PCGs, tRNAs and rRNAs) were extracted into a new folder and a phylogenetic tree was generated for each mt gene using Geneious tree builder (version 8.1). Tamura-Nei genetic distance model and the Neighbor-Joining tree build method were applied. Also, bootstrap resampling using 100 replicates was performed to assess node support values of the trees. Finally, a 50% majority-rule consensus tree was built, which included those clades present in the majority

of the trees of the original set of the resampled trees (Biomatters 2015). For determining the evolutionary distances of the six potato psyllid populations, to infer the true phylogeny, the Tamura-Nei distance model was used since Geneious only offers two options: Jukes-Cantor and Tamura-Nei. Unlike Jukes-Cantor, Tamura-Nei takes into account unequal frequencies in all bases. Neighbor-joining method, is the most widely algorithm used for developing quick trees and from Geneious options: neighbor-joining and UPGMA, neighbor-joining is the most suitable for the data. The bootstrap analysis was applied to evaluate the strength of clade support in the generated phylogenetic tree. In addition, a double cluster dendrogram was performed using the patristic distances, sum of the lengths of the branches that link two nodes in a tree and summarize the genetic change (Fourment and Gibss 2006), given by each single mt gene inference tree and using Washington (chosen arbitrary) as the population reference. The double cluster dendrogram was performed by writing a script that included the patristic distances of the populations and the command was run using R, applying Heatmap2 (gplots package) and color brewer.

Results

A total of 37 phylogenetic trees (Figure 3.1-3.3), corresponding to the 37 genes of the mitochondrial genomes of the different potato psyllid populations: central biotype (Texas and Nebraska), western biotype (California and Washington), northwestern haplotype and southwestern haplotype, were built. *P. sinica* (the wolfberry psyllid) was used as outgroup, since this organism was hypothesized to be closely related enough to the potato psyllid, but not as closely related as the other potato psyllid populations are to each other. This close relationship is supported by both *P. sinica* and *B. cockerelli* being important pests of solanaceous plants (but different genus), and belong to the same suborder of Hemiptera (*Sternorryncha*), but to different families.

In general, the populations' relationships were positioned on the trees as follows: northwestern, California, Nebraska, southwestern, and cluster together Texas and Washington

(Figures 3.1 and 3.3). These relationships were seen in trees based on the majority of PCGs and tRNAs. Another relationship seen in trees based on tRNAs was: California, Nebraska, northwestern, southwestern, Texas and Washington (Figure 3.3). These population arrangements were supported with high bootstrap values (93-100). The only genes that were consistent with the relationship seen in the literature was atp6, cox1 and cox3 and rrnS (Figures 3.1 and 3.2). However, the relationships showed moderate low (50-63) bootstrapping support.

An overall view of how all the genes related the potato psyllid populations among each other was seen in the double cluster dendrogram (Figure 3.4). In the dendrogram, the genotype differences of northwestern from the other populations are supported. In addition, it can be seen how PCGs are more susceptible to variations, while tRNAs and rRNAs seem to be more conserve. Also, in the dendrogram, northwestern is by itself, Nebraska and California are clustered together, and Southwestern is clustered with Texas and Washington.

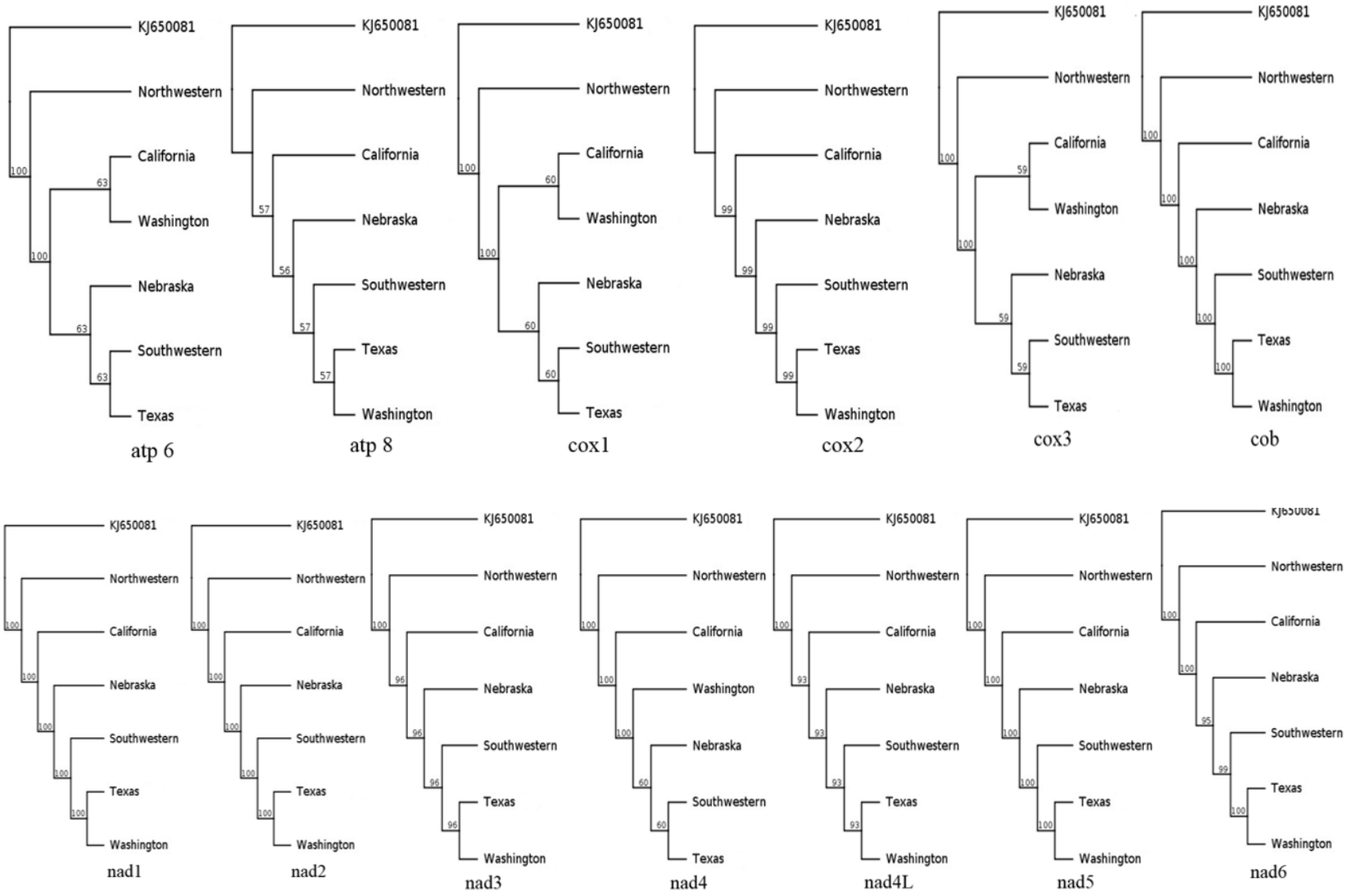

KJ650081
Northwestern
California
Washington
Nebraska
Southwestern
Texas
atp 6
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
atp 8
KJ650081
Northwestern
California
Washington
Nebraska
Southwestern
Texas
cox1
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
cox2
KJ650081
Northwestern
California
Washington
Nebraska
Southwestern
Texas
cox3
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
cob
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
nad1
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
nad2
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
nad3
KJ650081
Northwestern
California
Washington
Nebraska
Southwestern
Texas
nad4
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
nad4L
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
nad5
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
nad6

Figure 3.1. Phylogenetic tree based on the mitochondrial protein coding genes. Phylogenetic inference of the different potato psyllids (B. cockerelli Sulc) populations: Central biotype (Texas and Nebraska), Western biotype (California and Washington), Northwestern and Southwestern haplotype, and Paratrioza sinica (NCBI accession name KJ650081), the wolfberry psyllid, used as outgroup, clustering them by one of the protein coding genes of the mt genome: COI-III (cox1-3), Cytb (cob), ND1-6 (nad1-6), ND4L (nad4L), ATP6, and ATP8. The trees were built using: Geneious Tree Builder (Geneious version 8.1), applying the Tamura-Nei distance model, Neighbor-Joining method and Bootstrapping with 50% support threshold (Biomatters 2015). Bootstrap values of 95% or greater are considered statistically significant and give support to a clade.

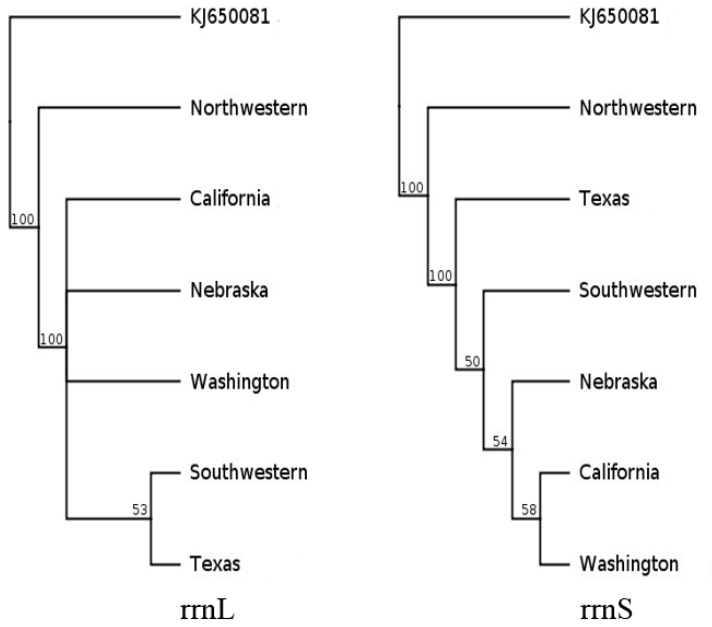
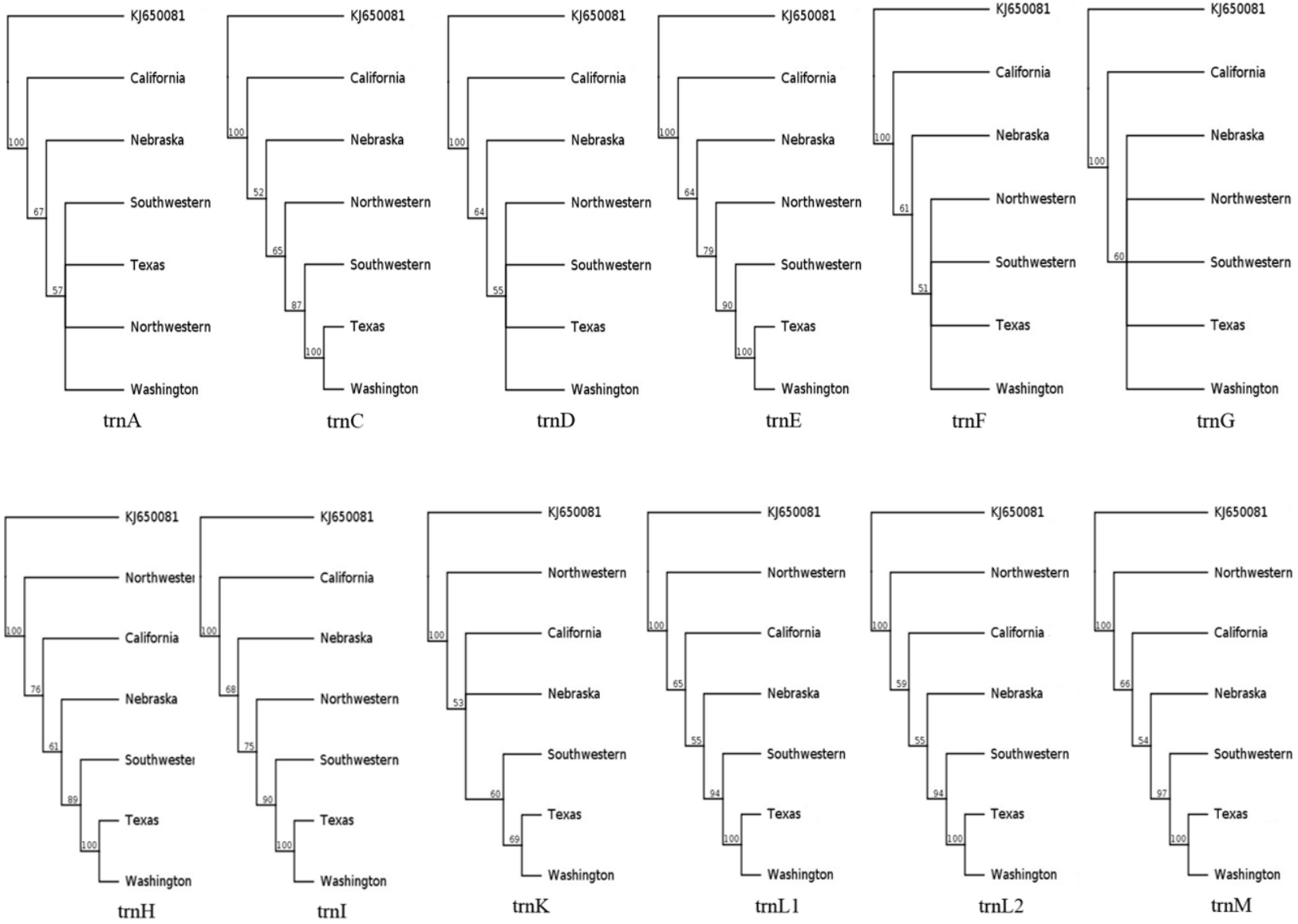


Figure 3.2. Phylogenetic tree based on the mitochondrial ribosomal RNAs. Phylogenetic inference of the different potato psyllids (B. cockerelli Sulc) populations: Central biotype (Texas and Nebraska), Western biotype (California and Washington), Northwestern and Southwestern haplotype, and Paratrioza sinica (NCBI accession name KJ650081), the wolfberry psyllid, used as outgroup, clustering them by one of the rRNA genes of the mt genome: : rrnL and rrnS. The trees were built using: Geneious Tree Builder (Geneious version 8.1), applying the Tamura-Nei distance model, Neighbor-Joining method and Bootstrapping with 50% support threshold (Biomatters 2015). Bootstrap values of 95% or greater are considered statistically significant and give support to a clade.


KJ650081
California
Nebraska
Southwestern
Texas
Northwestern
Washington
trnA
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnC
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnD
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnE
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnF
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnG
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnH
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnI
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnK
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnL1
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnL2
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnM

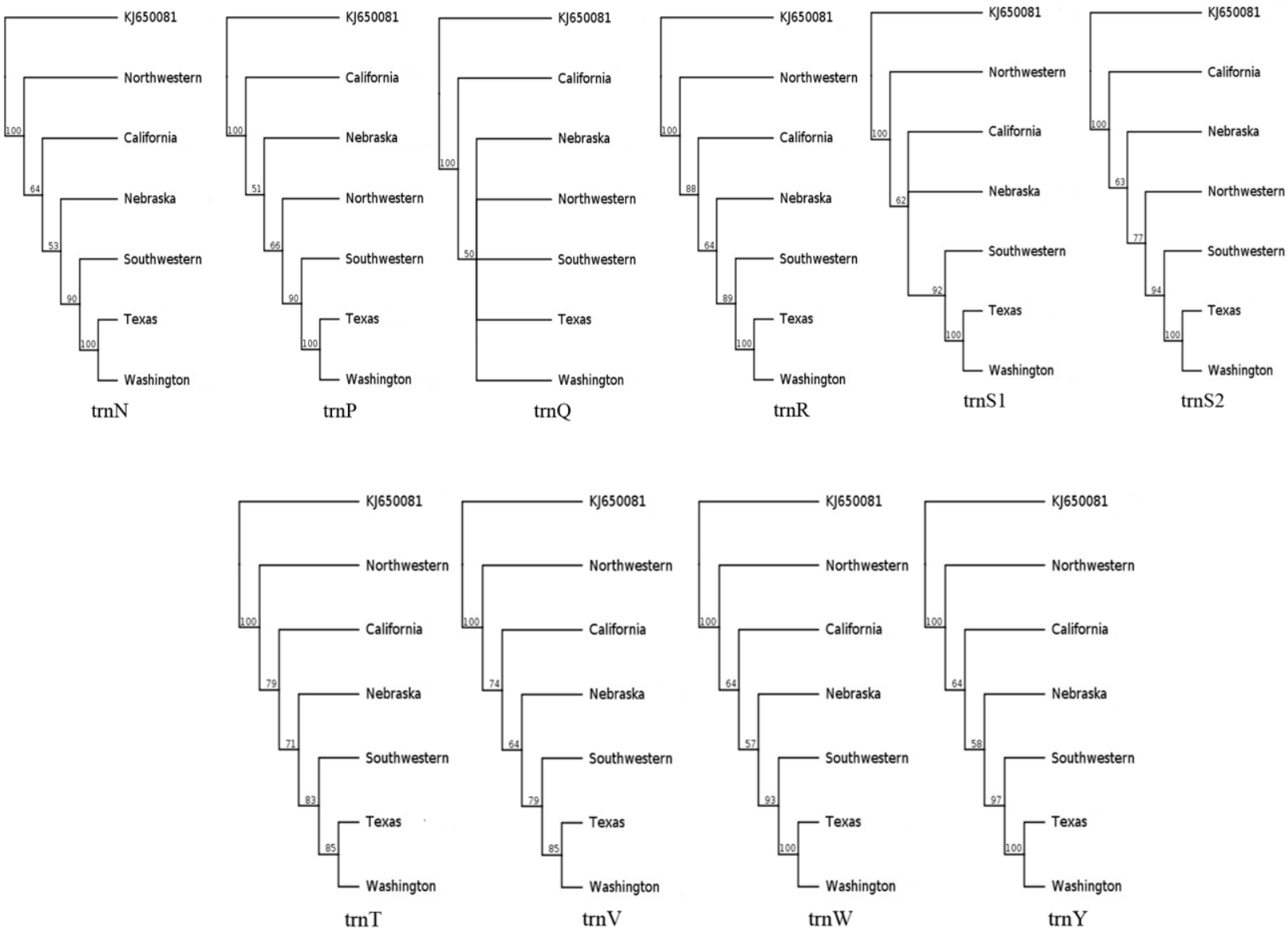

KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnN
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnP
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnQ
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnR
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnS1
KJ650081
California
Nebraska
Northwestern
Southwestern
Texas
Washington
trnS2
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnT
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnV
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnW
KJ650081
Northwestern
California
Nebraska
Southwestern
Texas
Washington
trnY

Figure 3.3. Phylogenetic tree based on the mitochondrial transfer RNAs. The phylogenetic inference of the different potato psyllids (B. cockerelli Sulc) populations: Central biotype (Texas and Nebraska), Western biotype (California and Washington), Northwestern and Southwestern haplotype, and Paratrioza sinica (NCBI accession name KJ650081), the wolfberry psyllid, used as outgroup, clustering them by one the tRNA genes of the mt genome. The trees were built using: Geneious Tree Builder (Geneious version 8.1), applying the Tamura-Nei distance model, Neighbor-Joining method and Bootstrapping with 50% support threshold (Biomatters 2015). Bootstrap values of 95% or greater are considered statistically significant and give support to a clade.

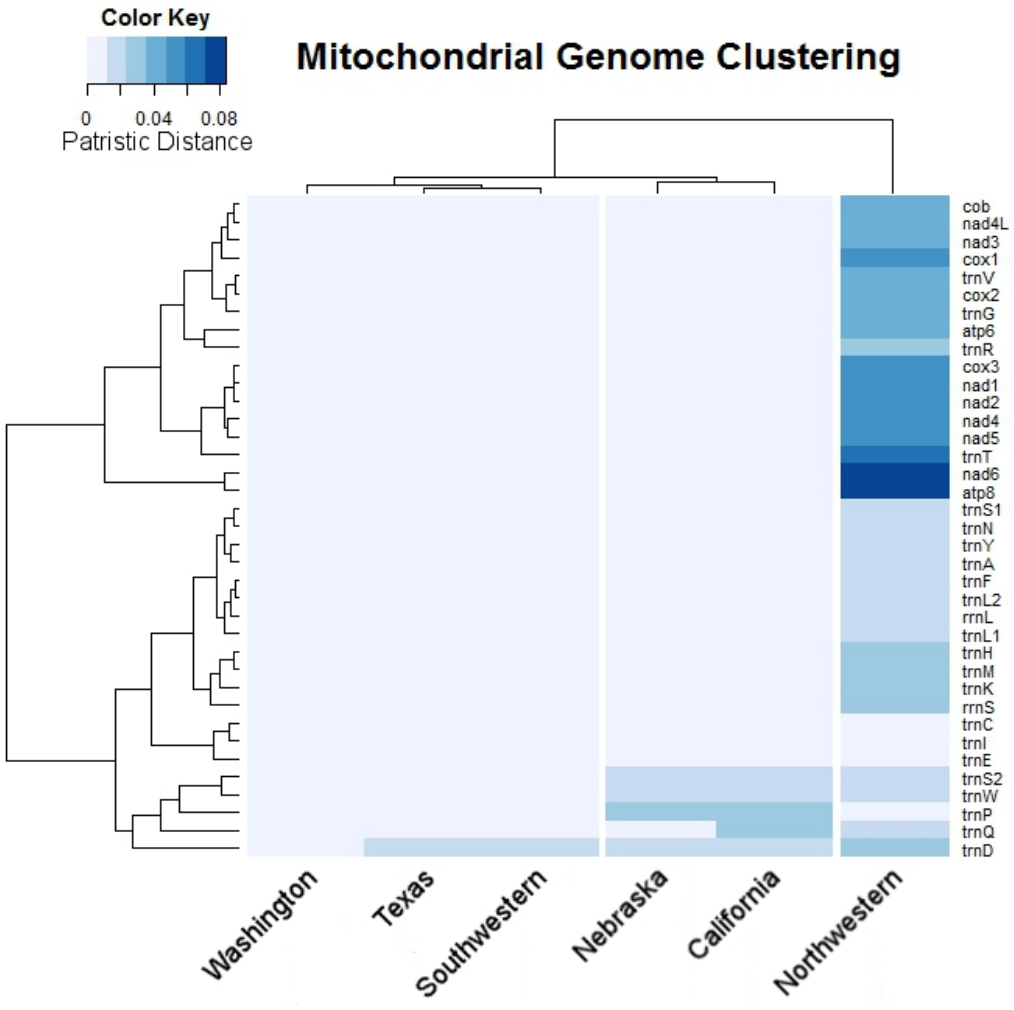


Figure 3.4. Double cluster dendrogram showing the genotypic similarity among B. cockerelli populations from Washington, Texas, Southwestern Nebraska, California and Northwestern (bottom). The darker the color of the genes, the further away genetically populations are from each other and viceversa. The dendrogram is based on the patristic distances, sum of the lengths of the branches that link two nodes in a tree and summarize the genetic change (Fourment and Gibss 2006), taken from the phylogenetic trees of the 37 mitochondrial genes (right), and Washington population was chosen, randomly, as the reference. The double cluster dendrogram was performed using R.

Discussion

A phylogenetic analysis of the 37 genes (2 rRNAs, 13 PCGS and 22 tRNAs) of the mitochondrial genome was performed to infer how each gene associated each potato psyllid population: Texas, Nebraska, California, Washington, southwestern and northwestern, among each other. Only the PCGs: atp6, cox1 (COI) and cox3 (COIII) cluster the different potato psyllid populations as seen using the whole mt genome: southwestern haplotype and central biotype (Texas and Nebraska) as one clade, as a second clade western biotype (California and Washington), and northwestern haplotype isolated from the others. COI output was not surprising, since it has been the gene marker in DNA barcoding widely used for species identification and delimitation (Cameron 2014), and the gene used for identifying and determining genotypic differences among the potato psyllid biotypes and haplotypes. Regarding the other genes, COIII and the ATPase genes have not been used as often as other PCGs for phylogenetic studies; however, these genes present amino acid sites that are free to vary and these unconstrained amino acid positions can provide useful information for studying relationships (Simon et al. 1994). In addition to the expected trees, among the majority of mt gene trees obtained, there is a clade that groups Texas and Washington together and is significantly supported by bootstrap values. In contrast, the expected clades between biotypes, in the expected trees, were moderately supported by these same values.

The nonparametric bootstrap is the most widely used method for evaluating nodal support in phylogenetic trees (Simon et al. 1996). Therefore, bootstrap values of 95% or greater are considered statistically significant and give support to a clade; while nodes with less than 5% of bootstrap estimates can be rejected (Soltis and Soltis 2003). Moreover, the interpretation of the bootstrap values can be opened to discussion and give a biased estimate of accuracy if the values are understood incorrectly. Bootstrap replicates are generated from the observed data, rather than the truth; thus, the value provides a confidence interval that contains the phylogeny that would be estimated from repeated sampling of many characters from the underlying set of all characters,

not the true phylogeny (Soltis and Soltis 2003; Simon *et al*. 1996). Hence, the high bootstrap values that support Texas and Washington relationship are nothing more than a consequence of the bootstrapping assessing method. As it is seen in the double cluster dendrogram (Figure 3.4), the lighter areas (the share genetic regions) are more abundant than the darker areas (the genetically different regions), which means that bootstrap replicates were generated from the most abundant observed data, all share genetic regions, and the proportion of this replicate data set the value. With that said, we can be confident of the relationship and how southwestern haplotype and central biotype cluster together and apart from the western biotype (California and Washington). Moreover, the double cluster dendrogram does shows how genotypically different is northwestern haplotype from the other populations and how genes such as the protein coding genes (darker color), are more prone to evolve because of the silent sites present on these genes that allow amino acid substitutions (Simon et al. 1994) than tRNA and rRNA genes that are more conserved (lighter color).

The faster rates of evolution in mt genes have been compared to the higher rates in which transition mutations occur. In comparison with the most rapidly evolving nuclear genes, it has been estimated, empirically, that synonymous substitutions of mt genes accumulate 1.7-3.4 times faster; and factors influencing these rates include: thermal adaptations, mitochondrial-nuclear interactions, and, in arthropods, infection with *Wolbachia* (Simon et al. 2006). Depending on their structure and function, rRNA, protein and tRNA genes evolve distinctly. Ribosomal genes are the most conserve genes of the mtDNA. Both the secondary structure and the ribosomal protein attachment sites of the molecule are highly conserved across all genomes, except for the ribosomal domains (peripheral sections) that are less conserved, and its rates of evolution vary depending on their functional constraints. In contrast, tRNA genes are less constrained by structure and function, evolving at a higher rate but slower than mt protein coding genes. Moreover, the triplet code nucleotide sequence of PCGs because of its coding function, possess strong constraints, in general, in the first and all second codon positions. The third position and

some first positions of the triplet are less restricted, and evolve at higher rates, either by silent substitutions or by amino acid replacement; but, as seen in rRNAs and tRNAs genes, the evolutionary rate will vary within protein coding genes in relation to its structural and functional constraints. Therefore, for studies of recently diverged taxa or intermediate levels of divergence PCGs are more reliable; but for studying deep levels of divergence, the highly conserved sites of rRNAs and tRNAs genes are better predictors of genetic divergence (Simon et al. 1994).

Single mt genes were used initially in insect phylogenetics because of the limited methods and protocols for sequencing nuclear genes and whole mt genomes. Since the beginning, individual mt genes have been used to answer the same inquiries when using whole mt genomes: phylogenetic relationships, population genetics and evolutionary history (Cameron 2014). Also, there has been a debate, whether to include all 37 genes of the mt genome when performing phylogenetic analysis. The animal mt genome contains 13 protein-coding genes that show different rates and patterns of nucleotide substitution within and between taxonomic groups (Bergmann et al. 2013), and in practice, almost all studies have included these 13 PCGs. The use of the rRNAs and tRNAs has been more variable because they are highly conserved, but in general, there is no justification to not include them in the analysis (Cameron 2014). In the case of the potato psyllid, data shows that the use of a single mt gene with high rate of nucleotide substitution such as COI gene can accurately determine phylogenetic relationships among the different populations. Faster evolving mt genes, such as COI, experienced more multiple substitutions and provide more resolving power for phylogeny of closely related taxa (Simon et al. 2006), power that has been noted with the determination of the potato psyllids biotypes and haplotypes (Liu et al. 2006; Chapman et al. 2012; Swisher et al. 2012). Whereas other genes like CytB, NAD1 and the large rRNA subunit, have not been effective for showing variations among potato psyllid populations (Powell et al. 2012). Nonetheless, for populations' genetics studies, the whole 37 genes should be included as they provide a more define and truly genotypic differentiation.

References

Ajmal Ali, M., G. Gyulai, N. Hidvégi, B. Kerti, F. M. A. Al Hemaid, A. K. Pandey, and J. Lee. 2014. The changing epitome of species identification – DNA barcoding. Saudi Journal of Biological Sciencies, 21: 204-231.

Bergmann, T., J. Rach, S. Damm, R. Desalle, B. Schierwater, and H. Hadrys. 2013. The potential of distance-based thresholds and character-based DNA barcoding for defining problematic taxonomic entities by CO1 and ND1. Molecular Ecology Resources 13: 1069-1081.

Biomatters. 2015. Geneious v8.1. www.geneious.com

Buchman, J., T. W. Fischer, V. G. Sengoda and J. E. Munyaneza. 2012. Zebra Chip Progression: From Inoculation of Potato Plants with Liberibacter to Development of Disease Symptoms in Tubers. Am. J. Pot Res 89:159-168.

Buchman, J. L., V. G. Sengoda and J. E. Munyaneza. 2011. Vector Transmission Efficiency of Liberibacter by *Bactericera cockerelli* (Hemiptera: Triozidae) in Zebra Chip Potato Disease: Effects of Psyllid Life Stage and Inoculation Access Period. J. Econ. Entomol. 104(5): 1486-1495.

Butler, C. and J. Trumble. 2013. Potato Psyllid. Center for Invasive Species Research (CISR) of the University of California, Riverside. Retrieved from: http://cisr.ucr.edu/potato_psyllid.html

Cameron, S. L. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu. Rev. Entomol. 59: 95-117.

Casiraghi, M., M. Labra, E. Ferri, A. Galimberti, and F. De Mattia. 2010. DNA barcoding: a six question tour to improve user's awareness about the method. Briefings in Bioinformatics, 2(4): 440-453.

Chai, H-N., Y-Z. Du, and B-P Zhai. 2012. Characterization of the Complete Mitochondrial Genomes of *Cnaphalocrocis medinalis* and *Chilo suppressalis* (Lepidoptera: Pyralidae). Int. J. Biol. Sci. 8(4): 561-579.

Chan, A., L-P. Chiang, H. C. Hapuarachchi, C-H. Tan, S-C. Pang, R. Lee, K-S. Lee, L-C Ng, and S-G. Lam-Phua. 2014. Dna barcoding: completing morphological identification of mosquito species in Singapore. Parasites and Vectors, 7: 569 doi:10.1186/s13071-014-0569-4.

Chapman, R. I., J. F. Macias-Velasco, A. P. Arp, and Blake R. Bextine. 2012. Using quantitative real time PCR melt curve analysis of partial CO1 sequence for rapid biotype differentiation of *Bactericera cockerelli* (Hemiptera: Triozidae). Southwestern entomologist, 37(4): 2012.

Collins, R. A., and R. H. Cruickshank. 2013. The seven deadly sins of DNA barcoding. Molecular Ecology Resources, 13: 969-975.

Friedrich, M., and N. Muqim. 2003. Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castanaeum*. Molecular Phylogenetics and Evolution 26:502-512.

Fourment, M., and M. J. Gibbs. 2006. PATRISTIC: a program for calculatin patristic distances and graphically comparing the components of genetic change. BMC Evol Biol. 6:1 doi: 10.1186/1471-2148-6-1

Golick, D. A., T. M. Heng-Moss, A. L. Steckelberg, D. W. Brooks, L. G. Higley, and D. Fowler. 2013. Using web-based key character and classification instruction for teaching undergraduate students insect identification. J Sci Educ Technol, 22: 509-521.

Janse, J. D. 2012. Bacterial disease that may or do emerge, with (possible) Economic damage for Europe and the Mediterranean Basin: notes on epidemiology, risks, prevention and management on first occurrence. Journal of Plant Pathology, 94(4): S4.5- S4.29.

Lee, W., and S-I. Akimoto. 2015. Development of new barcoding loci in gall-forming aphids (Eriosomatinae: Eriosomatini): comparing three mitochondrial genes, ATP6, ATP8, and COI. Journal of Asia-Pacific Entomology 18: 267-275.

Li, H., J. Gao, H, Liu, H. Liu, A. Liang, X. Zhou, and W. Cai, 2011. The Architecture and Complete Sequence of Mitochondrial Genome of an Assassin Bug *Agriosphodurs dohrni* (Hemiptera: Reduviidae). Int. J. Biol. Sci, 7(6): 792-804.

Liu, D., J. T. Trumble, and R. Stouthamer. 2006. Genetic differentiation between eastern populations and recent introductions of potato psyllid (*Bactericera* cockerelli) into western North America. Entomologia Experimentalis et Applicata 118: 177-183.

Nachappa, P., J. Levy and C. Tamborindeguy. 2012. Transcriptome analyses of *Bactericera cockerelli* adults in response to "*Candidatus* Liberibacter solanacearum" infection. Mol Genet Genomics 287: 803-817.

Meeyen, K., P. Nanork Sopaladawan, and P. Pramual. 2014. Population structure, population history and DNA barcoding of fruit fly *Bactrocera latifrons* (Hendel) (Diptera: Ttephritidae). Entomological Science, 17: 219-230.

Munyaneza, J. E. 2012. Zebra Chip Disease of Potato: Biology, Epidemiology, and Management. Am. J. Pot Res 89: 329-350.

Munyaneza, J. 2010. Psyllids as vectors of Emerging Bacterial Diseases of Annual Crops. Southwestern Entomologists, 35(3): 471-477.

Munyaneza, J.E., J.M. Crosslin and J. E. Upton. 2007. Association of *Bactericera* cockerelli (Homoptera: Psyllidae) with "Zebra Chip", a New Potato Disease in Southwestern United States and Mexico. Journal of Economic Entomology 100 (3): 656-663.

Powell, C. M., J. King, and B. R. Bextine. 2012. Cytochrome B Sequences of Potato Psyllids, *Bactericera cockerelli* (Sulc), from North and Central America. Southwestern Entomologist, 37(4): 521-523.

Ramírez-Davila, J.F., E. Porcayo-Camargo, J. R. Sánchez-Pale, and L.M. Vásquez-García. 2012. Spatial Behavior Comparison of *Bactericera cockerelli* Sulc. (Hemiptera: Triozidae) in Mexico. Neotro Entomol 41: 9-16.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87(6): 651-701.

Simon, C., T. R. Buckley, F. Frati, J. B. Stewart, and A. T. Beckenbach. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved

polymerase chain reaction primers for animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 37: 545-579.

Soltis, P. S., and D. E. Soltis. 2003. Applying the bootstrap in phylogeny reconstruction. Statistical Science, 18(2): 256-267.

Swisher, K. D., J. E. Munyaneza, and J. M. Crosslin. 2012. High Resolution Melting Analysis of the Cytochrome Oxidase I Gene Identifies Three Haplotypes of the Potato Psyllid in the United States. Environmental Entomology, 41(4): 1019-1028.

Chapter 4

Identification of Insecticide Resistance Genes in Potato Psyllid and its Variations among Potato Psyllid Populations

Abstract

Pests in agriculture, such as the potato psyllid, have been managed through the use of insecticides to decrease the damage that they cause through population reduction. Through time, the use of these chemicals has promoted the development of resistance for survivorship. Multiple genes are known to be responsible for this resistance either by mutations in target proteins, or by the enhancement of detoxifying enzymes, such as esterases, cytochrome P450s (CYP450s), and glutathione. So far, studies of effects of insecticides on the behavior and fitness of the potato psyllid, have found that the use of certain insecticides have conferred resistance to potato psyllid populations. However, which genes are involved is unknown. In this study, the potato psyllid transcriptome data was examined for the search of genes that are involved in insecticide resistance. Genes such as GSTs, GST-E1, GST-O1, CYP450, CYP4G, acetylcholinesterase and carboxylesterase were selected, and PCR primers were designed using NCBI Primer-Blast. Genes that showed PCR amplification (GST-O1 and CYP4G) in potato psyllid individuals from central and western biotypes, and southwestern and northwestern haplotypes, were sequenced but only successful results were gotten for CYP4G gene. The resulting sequences from this gene were used for building a phylogenetic tree, in which the western biotype, central biotype and southwestern haplotype appeared to be most closely related, and divergent from this clade is the northwestern haplotype. Resistance genes, such as CYP4G can be another approach of

differentiating among potato psyllid populations, since it seems that these genes are differentially expressed.

Introduction

Insecticides have played an important role in managing species of insects that are major pests in agriculture (Broekgaarden et al. 2011). However, the use (and sometimes overuse) of insecticides has promoted the development of insect resistance through adaptation. The rate of resistance development within a population of insects depends on the levels of genetic variability within the population and is facilitated by the pressure to adapt. Thus, resistance is an evolutionary phenomenon, in which pre-adaptive individuals that have not been exposed to insecticides, carry an altered genome that results in one or more possible mechanisms for survival from the selection pressure of insecticides (Li et al. 2013). Multiple genes are responsible for the development of the mechanisms of insecticide resistance, and the two most common mechanisms are target-site modifications and enzyme-based resistance. Target-site modifications involve a mutation that causes the insecticide to not bind to its target. Enzyme-based resistance consist of the enhancement or modification of the activity of detoxifying enzymes, such as esterases, cytochrome P450s (CYP450s), and glutathione S-tranferases (GSTs), which in turn prevents the insecticide to reach its molecular target by increasing the metabolism of insecticides into less harmful substances to facilitate its excretion (Hardy 2014; Li et al. 2013).

Glutathion S-transferases are a large family of detoxifying enzymes encoded by a complex multi-gene family in insects (Ffrench-Constant 2013; Chen et al. 2003). GSTs are classified in several classes (zeta, theta, sigma, and omega) and are involved in fundamental roles of basic metabolism, such as metabolic and signaling pathways (Ffrench-Constant 2013; Ketterman et al. 2011; Chen et al. 2003). Also GSTs have the ability to confer insecticide resistance by conjugating glutathione to xenobiotic (drugs, pesticides, plant toxin, chemical carcinogens, mutagens and others) compounds with electrophilic centers, converting insecticides

(reactive lipophilic molecules) into water-soluble non-reactive conjugates that are easily excreted (Chen et al. 2003). GSTs have been associated with resistance to organophosphates, organochlorine, and pyrethroid insecticides. In addition, GTSs play two roles in insecticide resistance, the first one by binding and sequestering insecticides and the second by protecting against oxidative stress when this is a by-product of insecticidal toxicity, such as pyrethroids (Ffrench-Constant 2013; Ketterman et al. 2011). Besides GSTs, Cytochrome P450s (CYP450s) are also a large class of enzymes that perform critical biological functions such as the detoxification and/or activation of xenobiotics, as well as in the metabolism of endogenous compounds like hormones, fatty acids, and steroids (Liu et al. 2015). CYP450s are divided into various CYP families, based on the degree on amino acid sequence identity, which include: 4, 6, 9, 12, 15, 18 and 28; but only families 4, 6, and 9 have been frequently linked to insecticide metabolism and resistance. In insects, CYP450s are represented by around 48-164 genes (Tiwari et al. 2011), and perform important functions in insect growth, development, and reproduction, as well as in the development of resistance. Studies have revealed that multiple CYP450s genes are upregulated in individual resistant organisms, demonstrating elevated expression levels of P450 genes in resistant insects (Liu et al. 2015).

In North America, species of insects from the orders: Hemiptera, Diptera, and Hymenoptera are among the most economically important insect pests. In this matter, Hemiptera is represented by 25 families, including Psyllidae with 2 species: *Diaphorina citri* and *Bactericera cockerelli* sulc. (Hardy 2014), vectors of huanglobing disease of citrus and zebra chip (ZC) disease of potatoes. The potato psyllid, *Bactericera cockerelli* sulc., is a phloem feeder that transmits the phytopathogen, *Candidatus* Liberibacter solanacearum (Lso), which causes ZC disease in potatoes and affects other solanaceous. The insect can cause a direct damage through feeding by nymphs, which causes plants' yellowishing and underdevelopment known as the "psyllid yellows disease", which consequently affects tuber yield and quality (Munyaneza et al. 2007; Ramírez-Davila et al. 2012; Liefting et al. 2009). In addition, the potato psyllid adult can

cause an indirect damage by the transmission of phytoplasmas (purple top disease) or of Lso by both nymphs and adults (Ramírez-Davila et al. 2012; Buchman et al. 2011; Liefting et al. 2009). The potato psyllid has had a great economic impact on the potato industry because ZC reduces the quality and longevity of the crop (Butler et al. 2011). Infected plants present, besides chlorosis and other like disease symptoms, a pattern of brown stripes in the tuber that are more visible when tubers are processed to produce potato chips (Munyaneza et al. 2007). These brown stripes result in tuber necrosis, which constitutes a lethal condition to the plant (Wen et al. 2009). Additionally, ZC causes the abandonment of whole areas dedicated for potato growth (Gao et al. 2009).

Since 1915, *B. cockerelli* was recognized as a plant pest because of damages to the False Jerusalem Cherry (*Solanum capsicastrum*) in California (Butler and Trumble 2012). Ever since, sporadic outbreaks of the potato psyllid occurred in potatoes in different states of the United States (every 20 or 30 years), but it was until 2001 that these outbreaks became recurrent each year in the US, and spread to Canada, Guatemala, Honduras and New Zealand (Butler and Trumble 2013; Butler and Trumble 2012; Abdullah 2008). Pest management in the US have used insecticides to manage the populations of potato psyllid. Along the years, the chemicals used include oils, nicotine, pyrenthrum, zinc arsenite spray, calcium cyanide dusts, lime-sulfur, DDT, organophosphates (i.e. phorate, parathion, disulfoton, and demeton), carbamates, pyrethoid insecticides, imidacloprid, spinosad, Kaolin particle film, and pymetrozine (Butler and Trumble 2012).

The major classes of insecticides: organochlorines, carbamates, organophosphates and pyrethroids, target the central system of the insect by interfering the propagation of the nervous impulse at the synaptic cleft or along the taxon (Montella et al. 2012). Imidacloprid, a systemic neonicotinoid insecticide, has been the most insecticide used for managing ZC and for potato psyllid control. This insecticide acts as an agonist of acetylcholine binding to post-synaptic nicotinic acetylcholine receptors, disrupting feeding behaviors and causing tremors, convulsions

and death of insects (Butler et al. 2012). In potato psyllids, resistance to imidacloprid has been found in populations of potato psyllids of California (Liu and Trumble 2007). Moreover, several studies have confirmed that the presence of a pathogen influences insecticide resistance (Tiwari et al. 2011). Arp et al. (2014) and Hail (et al. 2012) have suggested that microbial communities, especially the secondary endosymbionts in the potato psyllid, can give resistance to insecticides. Resistance to insecticides involves molecular and therefore physiological changes in the insect, which allows the survivorship of the populations under insecticide conditions, but which genes or group of genes are involved in potato psyllid resistance mechanism is still unknown. This study aim to identify resistance genes (esterases, GSTs, and CYP450s) in potato psyllid and to determine if the expression of these genes are influenced by biotype, haplotype and/or location.

Materials and methods

Primer designing

The potato psyllid transcriptome data (Fisher *et al*. 2014) available in http://www.sohomoptera.org/ACPPoP/index.html, was downloaded and examined to find genes that are known to be involved in insecticide resistance, such as esterases (carboxylesterases and acetylcholinesterases), cytochrome P450s, and glutathione S-tranferases. Once the sequence for each resistance gene (GSTs, GST-E1, GST-O1, CYP450, CYP4G, acetylcholinesterase and carboxylesterase) was selected from the database, NCBI Primer-Blast was used to design PCR primers from resistance gene sequences of 700 bp to 2kb in length. The primer designing settings included a PCR product size of 100 – 1500 bp, primer melting temperatures in a range of 55-65°C, and Hemiptera was selected as the organism for primer pair specificity checking parameters.

DNA preparation and PCR amplification

Individual potato psyllids belonging to central and western biotype, and southwestern and northwestern haplotype were used for nucleic acid extractions using the CTAB (cetyltriethylammonium bromide) buffer protocol (Zhang et al. 1998). Prior to DNA amplification, DNA concentrations were measured using 2 µl of DNA in a NanoDrop 1000 Spectrophotometer (Thermo Scientific, Waltham, MA).

The resistance genes primers sets: GSTs, GST-E1, GST-O1, CYP450, CYP4G, acetylcholinesterase and carboxylesterase, were optimized by performing a temperature gradient using conventional PCR to determine the optimum annealing temperature of the primers in the temperature range 50-60°C. Each 25 µl reaction included 12.5µl AmpliTaq Gold 360 Master Mix (Applied Biosystems, Foster City, CA), 1µl GC enhancer (Applied Biosystems, Foster City, CA), 2µl of resistance gene primer forward and reverse, 6.5µl water, and 1 µl of extracted DNA, which corresponded to a Texas sample. The thermal profile was run as follows: cycle 1, 95°C for 10 min (1x), cycle 2, 95°C for 30s, 50-60°C for 30s, 72°C for 60s (25x), and cycle 3, 72°C for 7 min (1x). Gel electrophoresis was run using 1% and 2% agarose gels (depending on the amplicon size) stained with ethidium bromide, and visualize using a UVP BioDock It Imaging System (Upland, CA). Optimization of primer concentration (range from 0.5µl to 1.75 µl per reaction) was also done for the primers sets that show amplification of the desired amplicon.

The resistance genes primers that worked, GST-O1 and CYP4G (cytochrome P450 family 4G), were run for all four potato psyllid populations. The 25 µl reaction was modified from the temperature gradient assay by 0.5µl of resistance gene primer forward and reverse, 6.5µl water, and 4 µl of extracted DNA. The thermal profile was run as follows: cycle 1, 95°C for 10 min (1x), cycle 2, 95°C for 30s, 50°C for 30s, 72°C for 60s (40x), and cycle 3, 72°C for 7 min (1x). Visualization of the PCR products was done by gel electrophoresis, as explained before.

Sequencing

To confirm the results of the PCR amplifications, PCR products from two individuals from each biotype and haplotype was purified for sequencing using Zymoclean Gel DNA Recovery Kit (Zymo Research Corp., Irvine, CA) following the provided protocol, and eluting the DNA with water. Sequencing was completed by Eurofins MWG Operon LLC DNA sequencing facility (Louisville, KY). The returned sequences were aligned using BioEdit (v7.2.5) and a consensus sequences was created for each population. Also, a tree was generated using Geneious tree builder (version 8.1), and the predicted *Diaphorina citri* (Asian citrus psyllid) CYP450 4G-like (NCBI reference sequence: XR_541492.1), was used as outgroup. The Tamura-Nei Distance Model and Neighbor-Joining Method with bootstrap resampling using 100 replicates (Biomatters 2014) was applied to assess node support.

Results

Seven sequences of different genes associated to insecticide resistance: GSTs (sequence ID BcWN_15254), GST-E1 (sequence ID BcWN_06156), GST-O1 (sequence ID BcWN_13635), CYP450 (sequence ID BcWN_06434), CYP4G (sequence ID BcWN_09364), acetylcholinesterase (sequence ID BcWN_06137) and carboxylesterase (sequence ID BcWN_07936), were used for designing primers. A maximum of five or minimum of three different primer sets, were ordered for each selected gene, depending on the number of primer return from the Primer-Blast. The primer set of GST-O1 (AL_GST-O1_F 5'-ACGGTGAGCTGCTGTAACTG -3'; and AL_GST-O1_R 5'-CATAGGCTGTAGCAAGCCCG -3') with an amplicon of 103 bp, and the primer set of CYP4G (CYP4G-F_AL 5'-GGGTAGGGCACAAACTCCTC -3'; and CYP4G-R_AL 5'- CGACTGCCACGACTACATGA -3') with an amplicon of 1,193 bp were successful at producing a clear band (Figure 4.1) for the four potato psyllid populations (Figure 4.2). In addition, with the resulting sequences a phylogenetic tree (Figure 4.3) was built using the consensus sequences obtained from the

resistance gene CYP4G for each potato psyllid population: central, western, southwestern, and northwestern. The predicted *Diaphorina citri* (asian citrus psyllid) CYP4G-like was used as outgroup, because there are not many species to choose from that this type of genes have been identified. Also, *D. citri* is closely related enough to the potato psyllid to not expect a direct relationship with the ingroup, which was supported with the BlAST search that resulted in a 79% identity to *D. citri* CYP450 4G-like. The gene GST-O1 did not gave any results from the sequencing.

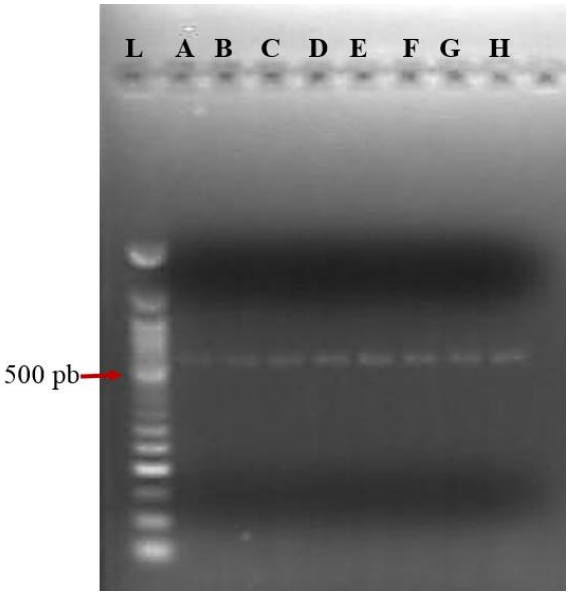


Figure 4.1. 1% agarose gel of the temperature gradient (50-60 °C) PCR for the resistance gene GST-O1. The primer pair designed was AL_GST-O1_F 5'- ACGGTGAGCTGCTGTAACTG -3'; and AL_GST-O1_R 5'-CATAGGCTGTAGCAAGCCCG -3'. The gel was stained with ethidium bromide. The initials on the gel refer to: L (ladder) = Quick-Load 50 pb DNA Ladder (New England Biolabs, Ipswich, MA); and letters A to H the different annealing temperatures that were used, A=50°C, B=50.8°C, C=52.1°C, D=53.9°C, E=56.4°C, F=58.3°C, G=59.5°C, and H=60°C.

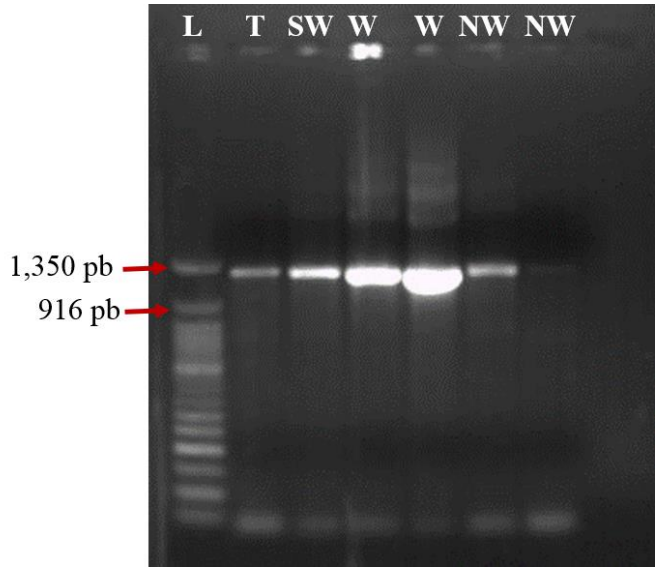


Figure 4.2. 1% agarose gel for the resistance gene CYP4G. The gel shows the bands obtained for the potato psyllid biotypes and haplotypes by PCR amplification of the resistance gene CYP4G primer pair designed: CYP4G-F_AL 5'-GGGTAGGGCACAAACTCCTC -3'; and CYP4G-R_AL 5'- CGACTGCCACGACTACATGA -3', with an amplicon of 1,193 bp, and stained with ethidium bromide. The initials on the gel refer to: L (ladder) = Quick-Load 50 pb DNA Ladder (New England Biolabs, Ipswich, MA); T (Texas) = central biotype, SW = southwestern haplotype, W = western biotype, and NW = northwestern haplotype.

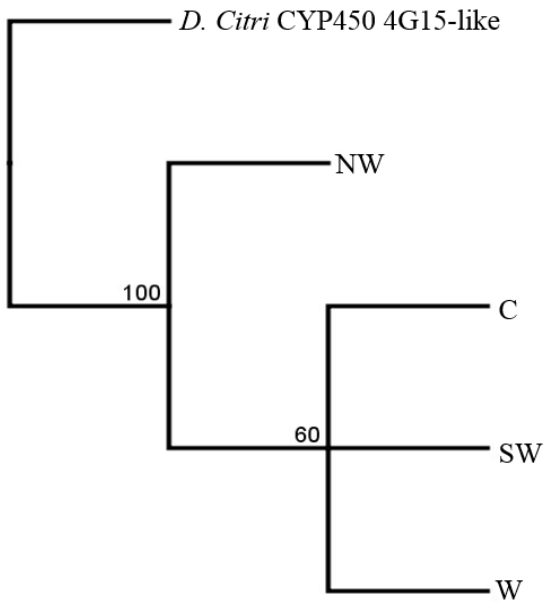


Figure 4.3. Phylogenetic inference of the potato psyllid (B. cockerelli Sulc) biotypes and haplotypes using insecticide resistance gene CYP4G. The potato psyllid biotype are: central (C) and western (W), and the potato psyllid haplotypes: northwestern (nw), and southwestern (sw). The outgroup chosen is the predicted D.citri (asian citrus psyllid) CYP450 4G-like. The tree was built using: Geneious Tree Builder (Geneious version 8.1), applying the Tamura-Nei distance model and Neighbor-Joining method, and Bootstrapping with 50% support threshold (Biomatters 2014).

Discussion

Insecticide resistance genes were identified in potato psyllid biotypes from central and western, and potato psyllid haplotypes from southwestern and northwestern. The genes identified were the detoxifying enzymes GST-O1 and the CYP family 4G (CYP4G). However, GST-O1 amplified a larger amplicon (approximately 550 bp), than the one expected (103 bp), and the resulting sequencing gave no results, which correlated to the amplification of the wrong amplicon. The latter can be likely because of incorrect primer design, primers are not amplifying the gene of interest, contamination or incorrect annealing temperature. Sequencing of the resulting amplicon was done to check if the primer was amplifying the target DNA, but no results came back. Also, as it is suggested, the temperature gradient PCR used for this gene was in a range of annealing temperatures that were 5°C above and below the primer melting temperature, but all the annealing temperatures amplified the same amplicon (Figure 4.1). Therefore, if desiring to continue working with this gene, lower annealing temperatures can be applied to check if there is any difference of amplicon size, or new sets of primers can be designed. Regarding CYP4G gene, both amplicon and sequencing results confirm the presence of this gene in the potato psyllid biotypes and haplotypes. Studies in related species, *Diaphorina citri*, have shown that CYP4G genes are involved in the metabolism of imidacloprid (Tiwari et al. 2011); and a study in midges (*Chironomus riparius*), revealed the potential of CYP4G as biomarkers for ecotoxicological studies (Martinez-Paz et al. 2012). Furthermore, the resulting negative data for all the other genes that were selected: GSTs, CYP450, acetylcholinesterase and carboxylesterase, might be due to the presence of large introns between the primer sites, which can lead to an incomplete formation of the desire amplicon, and resulting in no amplification. Also, GC contents that are too high can lead to nonspecific primer binding, which was not the case since primers were designed with a GC content between 40-60%. Thus, in silico PCR should be used to maximize the efficiency and selectivity of the designed primers that can help to reduce those negative PCR runs results.

The phylogenetic analysis built using CYP4G gene showed southwestern haplotype as well as western and central biotype as closely related. The western biotype includes populations from California, New Mexico up to Washington, Oregon and Idaho, and southwestern predominantly occurred in New Mexico and Southern Colorado. The central biotype occurs from eastern Mexico up to Texas, Kansas, Colorado, Nebraska, and Wyoming (Swisher et al. 2013). Since southwestern populations overlap with both central and western biotype, there is reason to believe that gene flow is occurring between the three of them and the genes to adapt to insecticide resistance have been carried out to either of the populations. Studies have found that potato psyllids from California populations have shown resistance to the insecticide imidacloprid (Liu and Trumble 2007). In addition, because of ZC chemical management, growers from Texas have stopped the use of the insecticide Monitor, an organophosphate, because of loss of efficacy after a few years of heavy use (Schreiber et al. n.d.), resulting in the development of insecticide resistance in the central biotype. Concerning the northwestern haplotype, is a population that geographically has been separated from the other potato psyllid populations. Until late 2011, ZC was documented for the first time in the Pacific Northwest, and little is known about managements in potatoes fields of this region (Schreiber et al. n.d.). Hence, it is reasonable that potato psyllids from this area do not present the same genotypic insecticide resistance variability of the other populations. Furthermore, the phylogeny obtained with the insecticide resistance gene CYP4G, despite the fact that cannot be accurate because the fast evolving rate of this type of genes, appeared to be phylogenetically informative in terms of chemical management within potato psyllid populations.

The selection pressure of insecticides have influenced populations to evolve and adapt, resulting in molecular changes that confer resistance to insecticides. In potato psyllids, resistance to insecticides has already been documented, and although CYP4G gene was present in all four populations, it seems that the molecular constitution of the gene can vary within location. Also, genes involved in insecticide resistance seem to be differentially expressed within the potato

psyllid populations, which can be associated to the presence of secondary endosymbionts. Arp et al. (2014) and Hail et al. (2012) have predicted that the purpose of potato psyllids having secondary endosymbionts might provide physiological alterations, such as resistance to insecticides. Similarly, surveys from Arp et al. 2014, indicated that microbial communities of the potato psyllid are primarily influenced by location, which in turn can enhance the expression of resistance genes. The latter has been documented in studies in mosquitoes, *Culex pipens*, which have shown that insecticide-resistant mosquitoes have much greater densities of *Wolbachia*, than insecticide-susceptible individuals (Echaubard et al. 2010). Nevertheless, this is only a preliminary study and for an accurate prediction of the expression of CYP4G genes and other related genes in potato psyllid populations, a large sample needs to be used.

References

Abdulla, N.M.M. 2008. Life history of the Potato Psyllid *Bactericera* cockerelli (Homoptera: Psyllidae) in controlled Environment agriculture in Arizona. African Journal of Agricultural Research 3(1): 060-067.

Arp, A., J. E. Munyaneza, J. M. Crosslin, J. Trumble, and B. Bextine. 2014. A Global Comparison of *Bactericera cockerelli* (Hemiptera: Triozidae) Microbial Communities. Environ. Entomol. 43(2): 344-353.

Biomatters. 2015. Geneious v8.1. www.geneious.com

Broekgaarden, C., T. A. L. Snoeren, M. Dicke, and B. Vosman. 2011. Exploiting natural variation to identity insect-resistance genes. Plant Biotechnology Journal, 9: 819-825.

Buchman, J. L., V. G. Sengoda, and J. E. Munyaneza. 2011. Vector Transmission Efficiency of Liberibacter by *Bactericera cockerelli* (Hemiptera: Triozidae) in Zebra Chip Potato Disease: Effects of Psyllid Life Stage and Inoculation Access Period. J. Econ. Entomol. 104(5): 1486-1495.

Butler, C. and J. Trumble. 2013. Potato Psyllid. Center for Invasive Species Research (CISR) of the University of California, Riverside. Retrieved from: http://cisr.ucr.edu/potato_psyllid.html

Butler, C. and J. T. Trumble. 2012. The potato psyllid, *Bactericera cockerelli* (Sulc) (Hemiptera: Triozidae): life history, relationship to plant diseases, and management strategies. Terrestrial Arthropod Reviews, 5: 87-111.

Butler, C. D., G. P. Walker, and J. T. Trumble. 2012. Feeding disruption of potato psyllid, *Bactericera cockerelli*, by imidacloprid as measured by electrical penetration graphs. Entomologia Experimentalis et Applicata, 142: 247-257.

Chen, L., P. R. Hall, X. E. Zhou, H. Rnason, J. Hemingway, and E. J. Meehan. 2003. Structure of an insect δ-class glutathione S-transferase from a DDT-resistant strain of the malaria vector *Anopheles gambiae*. Acta Crystallograhica Section D: 2211-2217.

Echaubard, P., O. Duron, P. Agnew, C. Sidobre, V. Noël, M. Weill, and Y. Michalakis. 2010. Rapid evolution of *Wolbachia* density in insecticide resistant *Culex pipiens*. Heredity, 104: 15-19.

Gao, F.; J. Jifon, X. Yang, and T. Liu. 2009. Zebra chip disease incidence on potato is influenced by timing of potato psyllid infestation, but not by the host plants on which they were reared. Insect Science, 16:399-408.

Hail, D., S. E. Dowd, and B. Bextine. 2012. Identification and Location of Symbionts Associated with Potato Psyllid (*Bactericera cockerelli*) Lifestages. Environmental Entomology 41(1): 98-107.

Hardy, M. C. 2014. Resistance is not futile: it shapes insecticide discovery. Insects, 5: 227-242.

Ffrench-Constant, R. H. 2013. The molecular genetics of insecticide resistance.
Genetics, 194: 807-815.

Ketterman, A. J., C. Saisawang, and J. Wongsantichon. 2011. Insect glutathione transferases. Drug Metabolism Reviews, 43(2): 253-265.

Li, M., W. R. Reid, L. Zhang, J. G. Scott, X. Gao, M. Kristensen, and N. Liu. 2013. A whole transcriptomal linkage analysis of gene co-regulation in insecticide resistant house flies, *Musca domestica*. BMC Genomics, 14: 803.

Liefting, L. W., P. W. Sutherland, L. I. Ward, K. L Paice, B. S. Weir, and G. R. G. Clover. 2009. A new "Candidatus Liberibacter" species associated with diseases of Solanaceous crops. Plant Disease, 93 (3): 208-214.

Liu, N., M. Li, Y. Gong, F. Liu, and T. Li. 2015. Cytochrome P450s – Their expression, regulation, and role in insecticide resistance. Pesticide Biochemistry and Physiology, doi: 10.1016/j.pestbp.2015.01.006

Liu, D. and J. T. Trumble. 2007. Comparative fitness of invasive and native populations of the potato psyllid (*Bactericera cockerelli*). Entomologia Experimentalis et Applicata, 123: 35-42.

Martinez-Paz, P., M. Morales, J. L. Martinez-Guitarte, and G. Morcillo. 2012. Characterization of a cytochrome P450 gene (CYP4G) and modulation under different exposures to xenobiotics (tributyltin, noylphenol, bisphenol A) in *Chironomus riparius* aquatic larvae. Comparative Biochemistry and Physiology, part C 155: 333-343.

Montella, I. R., R. Schama, and D. Valle. 2012. The classification of esterases: an important gene family involved in insecticide resistance – A review. Mem Inst Oswaldo Cruz, 107 (4): 437-449.

Munyaneza, J.E., J.M. Crosslin, and J. E. Upton. 2007. Association of *Bactericera* cockerelli (Homoptera: Psyllidae) with "Zebra Chip", a New Potato Disease in Southwestern United States and Mexico. Journal of Economic Entomology 100 (3): 656-663.

Ramírez-Davila, J.F., E. Porcayo-Camargo, J. R. Sánchez-Pale and L.M. Vásquez-García. 2012. Spatial Behavior Comparison of *Bactericera cockerelli* Sulc. (Hemiptera: Triozidae) in Mexico. Neotro Entomol 41: 9-16.

Schreiber, A., A. Jensen, and S. Rondon. n.d. Biology and management of Potato Psyllid in Pacific Northwest Potatoes. Zebra Chip IPM publication, Oregon Potato Commission.

Tiwari, S., A. D. Gondhalekar, R. S. Mann, M. E. Scharf, and L. L. Stelinski. 2011. Characterization of five CYP4 genes from Asian citrus psyllid and their expression levels in *Candidatus* Liberibacter asiaticus-infected and uninfected psyllids. Insect Molecular Biology, 20(6): 733-744.

Wen, A., I. Malik, V. Y. Alvarado, J. S. Pasche, X. Wang, H. B. Scholthof, T. E. Mirkov, C. M. Rush, and N. C. Gudmestad. 2009. Detection, distribution, and genetic variability of "*Candidatus* Liberibacter" species Associated with Zebra Complex disease of potato in North America. Plant Disease, 93(11):1102-1115.

Chapter 5

Conclusions and Future Research

Population studies of the potato psyllid have relied on the mitochondrial gene cytochrome c oxidase subunit 1 (COI) to identify and characterize variations among potato psyllid populations of the United States. Through the use of COI as a genetic marker, COI melt temperature and high resolution melting analysis, and sequencing data of this gene, the identification of biotypes (central and western) and haplotypes (northwestern and southwestern) of the potato psyllid were determined (Swisher *et al*. 2014; Swisher *et al.* 2013; Swisher *et al.* 2012; Chapman *et al.* 2012; Liu *et al*. 2006). This research focused on the use of all the genes that encompass the mitochondrial genome, as well as other genes (nuclear genes) for a better identification and understanding of the potato psyllid populations.

Complete mitochondrial genome sequences contain useful phylogenetic information, especially, when dealing with species such as the potato psyllid that are of economic importance and correct management approaches need to be taken. The development of molecular technology has facilitated the sequencing of whole mitochondrial genomes by reducing the costs and time. In this case, whole mitochondrial genome, instead of one single gene, improve the understanding of the genotypic relationships among the different potato psyllid populations by correctly defining a possible new biotype: northwestern, and removing the haplotype tag to the southwestern population. Also, the availability of more phylogenetic information allows to compare in itself, resulting in different approaches that can be taken into account when trying to characterize populations. For instance, single mitochondrial genes with highly susceptibility to variation as

COI, COIII and ATP6, can be applied to answer phylogenetic questions that do not require a complex background to be answered.

Moreover, in phylogenetic studies it has been recommended to combine both nuclear and mitochondrial data, since incongruences can be presented when working with only one of the genomes, and there is always a better outcome of important aspects of species histories when complete information is available (Simon *et al*. 2006). Thus, the identification of insecticide resistance genes in the potato psyllid, not only is another means of looking at variation among and within populations, but also a method for studying gene-environment interactions; as well as, developing efficient resistance management strategies since most of the zebra chip disease control involves the use of chemicals.

In addition, whole mitochondrial genomes are a powerful and useful data source that can be applied not only to the species of interest, but also for phylogenetic and comparative genomic analysis between taxa. For agricultural decisions, knowing how species that are vectors of pathogens, such as the potato psyllid, relate to others, can help to determine if the ones that are not considered a harm, can become in one. Therefore, given that the impact of zebra chip disease is a large enough concern, and since both vector and causal agent are expanding their geographic distribution north and south from the United States. It is crucial that potato psyllids from Central America are included in this type of studies to help in the understanding of the relationship between the vector and pathogen of the disease.

References

Chapman, R. I., J. F. Macias-Velasco, A. P. Arp, and Blake R. Bextine. 2012. Using quantitative real time PCR melt curve analysis of partial CO1 sequence for rapid biotype differentiation of *Bactericera cockerelli* (Hemiptera: Triozidae). Southwestern entomologist, 37(4): 2012.

Liu, D., J.T. Trumble, and R. Stouthamer. 2006. Genetic differentiation between eastern populations and recent introductions of potato psyllid (*Bactericera* cockerelli) into western North America. Entomologia Experimentalis et Applicata 118: 177-183.

Swisher, K. D., D. C. Henne, and J. M. Crosslin. 2014. Identification of a fourth haplotype of *Bactericera cockerelli* (Hemiptera: Triozidae) in the United States. J. insect Sci. 14(161): DOI: 10.1093/jisesa/ieu023

Swisher, K.D, J. E. Munyaneza, and J. M. Crosslin. 2013. Temporal and Spatial Analysis of Potato Psyllid Haplotypes in the United States. Environ. Entomol. 42(2): 381-393.

Swisher, K. D., J. E. Munyaneza, and J. M. Crosslin. 2012. High Resolution Melting Analysis of the Cytochrome Oxidase I Gene Identifies Three Haplotypes of the Potato Psyllid in the United States. Environmental Entomology, 41(4): 1019-1028.